Rethinking Regulations

The six papers in this series lays out a major paradigm shift in thinking about regulatory compliance and the best means for moving forward in a cost effective and efficient program monitoring system. These papers were published between 2024-2025.

Richard Fiene PhD

The Theory of Regulatory Compliance* and Its Implications for Regulatory Science

Richard Fiene PhD

Penn State Prevention Research Center

January 2024

The theory of regulatory compliance has appeared in a series of articles in the *Journal of Regulatory Science* and its spin off methodologies in other journals, *Child Care Quarterly, Child and Youth Forum, International Journal of Child Care and Education Policy, and Early Childhood Research Quarterly.* The theory has had a large impact on the human services, in particular the early care and education field. The purpose of this article is to reach a larger audience that may be representative of some of the other regulatory areas in the physical sciences, medical sciences and the economic sciences.

The organization of this article will first deal with the theory itself, explaining it in simple, non-mathematical terms and its implications for public policy and licensing decision making. Then we will delve into the implications and spin off methodologies of the theory, such as differential monitoring, risk assessment rule formulation, key predictor rules, the uncertainty-certainty matrix, ceiling effect, and dichotomization of skewed licensing data distributions.

Regulatory science is a relatively new science appearing on the scene in the past 20 years. Regulatory compliance and the licensing of programs, industries, etc. has been around for quite some time. The first licensing law was passed over 100 years ago governing orphanages in Pennsylvania. But as is clearly evident the science behind licensing and regulatory compliance lagged by many decades. Licensing grew at a slow pace in the human services during the

twentieth century and it was not until the late 1960's to early 1970's that human services began to really expand and grow in terms of the number of programs. Other industries grew in a corresponding way with most of the growth in later portions of the previous century. The pharmaceutical industry is a perfect example of this. In fact, regulatory science has really grown out of this need to regulate the pharmacological industry. The Food and Drug Administration is the leading federal agency in pursuing the expansion and dissemination of regulatory science with the establishment of national centers across the USA.

Let me provide some historical context to the theory and how it has evolved over the past several decades based upon empirical evidence. The original standard paradigm when it came to regulatory compliance and its relationship to program quality was that there is a linear relationship between the two components. As one goes up, the other goes up in a corresponding way. From a public policy standpoint this made a great deal of sense. Any licensing agency wants to see increased quality of services based upon their rules and regulations. I will only be addressing the human services, in particular early care and education programs, that is where all the research has been done. In the future, it will be necessary to determine if what is being described in the human services applies to industries outside of this domain.

The problem with the standard paradigm was that it was not based on empirical evidence but rather on expert opinion and anecdotal evidence, but there were no well-designed studies that looked at the relationship between regulatory compliance and program quality in any of the human services. Fast forward to the 1970's as the number of early care and education programs were increasing dramatically because of the influx of federal dollars as part of the

Great Society and the creation of Head Start and a major expansion of child care. It became clear that the standard paradigm which included doing case studies as their major means for data collection was not going to be a viable measurement strategy. This ushered in a new form of program monitoring and data collection called Instrument based Program Monitoring which utilized checklists, tools, and instruments for their data collection and licensing measurement.

Another thing that happened also in the movement from qualitative to quantitative measurement was that larger studies could be done to evaluate the relationship between regulatory compliance and program quality. Finally, there would be a chance to collect scientific data on this relationship and prove the linear relationship between regulatory compliance and program quality. When these studies were done, sure enough, when low levels of regulatory compliance which essentially means rule or regulatory violations are being found and comparing these data to the overall quality of the respective programs there was a direct linear relationship and that continued to be so right up to substantial regulatory compliance which means being 98-99% in compliance with all rules and regulations. However, then a very interesting change occurred in moving from substantial regulatory compliance to full (100%) regulatory compliance in which the respective programs did not follow the linear relationship and there was a plateauing or a ceiling effect in which it was difficult to distinguish the quality of programs that were in substantial vs full regulatory compliance. It was in some cases in subsequent studies (2010's) which replicated these initial studies in the 1970's where the relationship followed more of a diminishing returns type of curve. Not always but definitely a ceiling effect was always observed in the data.

These results obviously upset the proverbial public policy apple cart and the standard paradigm which was based upon a linear model and that licenses should only be issued to those programs that were in full regulatory compliance, no exceptions. The data did not support this claim nor the public policy. Substantial regulatory compliance was clearly demonstrating that these programs were providing the same level of quality care as those programs that were in full regulatory compliance and in some cases were doing an even better job of providing quality care. This is the major finding of the theory of regulatory compliance demonstrating these diminishing returns and/or ceiling effect and introduces substantial regulatory compliance as a licensing decision point rather than relying only on full 100% regulatory compliance. The original paradigm still holds in that regulatory compliance is very accurate in distinguishing between low and higher quality, but it is not as accurate when it comes to distinguishing quality at the substantial regulatory compliance and the full regulatory compliance levels.

The following figure/graphic (Figure 1) depicts the relationship between regulatory compliance levels and program quality scores. This graphic is a summary depiction of the various studies that have been completed starting in the 1970's through to the 2010's in looking at this relationship. The graphic also shows the relationship to several other concepts that will be addressed in this article, dealing with differential and integrative monitoring, key indicator predictor rules, risk assessment rules, nominal data, and dichotomization of data. All these additional concepts will be dealt with in the following sections of this article.

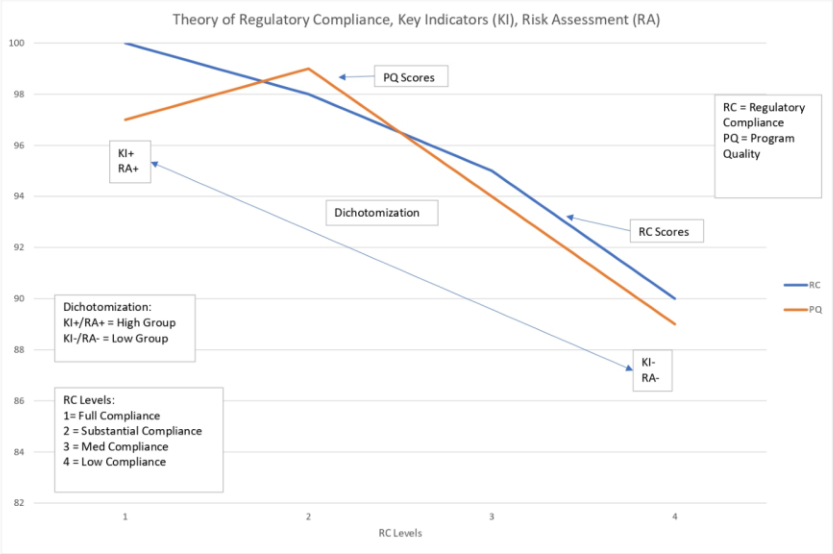


Figure 1: Theory of Regulatory Compliance

Let's turn our attention to some of the spin off methodologies and approaches from the theory of regulatory compliance. The first one to consider is differential monitoring because it is the most significant in altering the licensing landscape in how programs are monitored, reviewed, and inspected. Differential monitoring is about focused reviews rather than a one size fits all approach which again was predominant in the standard program monitoring paradigm. Because the theory of regulatory compliance introduced the importance of substantial regulatory compliance into the new and revised paradigm when it comes to program monitoring, it ushered in more targeted inspections or reviews which focused on key predictor rules or rules that placed clients at particular risk, more so than other rules and regulations. There was also part of this new paradigm the notion of reviewing programs less often but that was removed from the differential monitoring approach because all the research into program

monitoring indicated that just reviewing the program more frequently brought about more positive change in regulatory compliance and quality.

In Figure 2, the differential monitoring approach is depicted along with the definitions of each of the methodologies which are part of the approach.

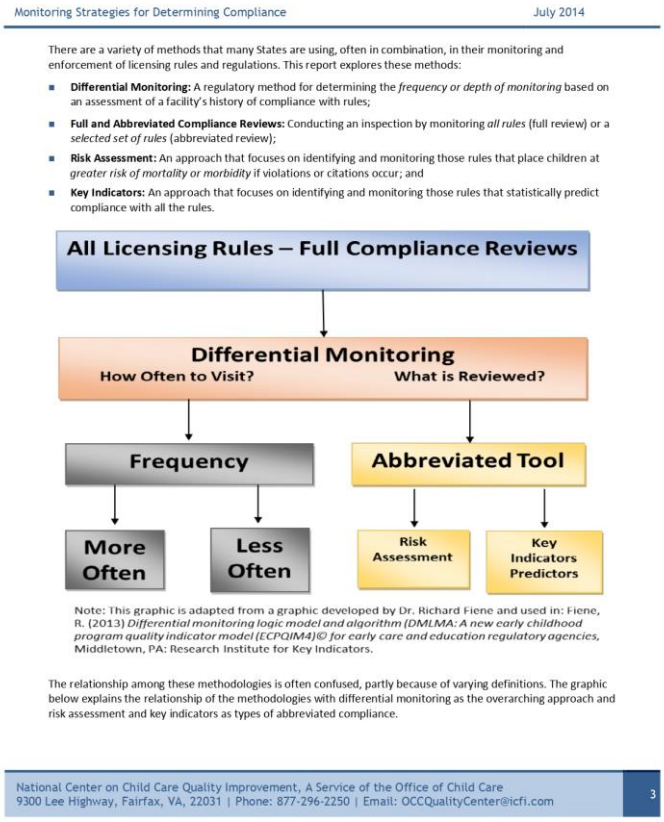

Monitoring Strategies for Determining Compliance
July 2014
There are a variety of methods that many States are using, often in combination, in their monitoring and enforcement of licensing rules and regulations. This report explores these methods:
Differential Monitoring: A regulatory method for determining the frequency or depth of monitoring based on an assessment of a facility's history of compliance with rules;
Full and Abbreviated Compliance Reviews: Conducting an inspection by monitoring all rules (full review) or a selected set of rules (abbreviated review);
Risk Assessment: An approach that focuses on identifying and monitoring those rules that place children at greater risk of mortality or morbidity if violations or citations occur; and
Key Indicators: An approach that focuses on identifying and monitoring those rules that statistically predict compliance with all the rules.
All Licensing Rules – Full Compliance Reviews
Differential Monitoring
How Often to Visit?
What is Reviewed?
Frequency
Abbreviated Tool
More Often
Less Often
Risk Assessment
Key Indicators Predictors
Note: This graphic is adapted from a graphic developed by Dr. Richard Fiene and used in: Fiene, R. (2013) Differential monitoring logic model and algorithm (DMLMA: A new early childhood program quality indicator model (ECPQIM4)© for early care and education regulatory agencies, Middletown, PA: Research Institute for Key Indicators.
The relationship among these methodologies is often confused, partly because of varying definitions. The graphic below explains the relationship of the methodologies with differential monitoring as the overarching approach and risk assessment and key indicators as types of abbreviated compliance.
National Center on Child Care Quality Improvement, A Service of the Office of Child Care
9300 Lee Highway, Fairfax, VA, 22031 | Phone: 877-296-2250 | Email: OCCQualityCenter@icfi.com
3

Figure 2: Differential Monitoring Approaches

Risk assessment is one of the methodologies which is part of the differential monitoring approach. It focuses on those specific rules and regulations which place clients/children are greatest risk of morbidity or mortality. These are the rules that deal with supervision,

hazardous materials being in locked cabinets, etc. Generally, jurisdictions/states/provinces can identify these rules through an empirical weighting approach where a Likert Scale is used to weight each rule or regulation on the basis of this morbidity and mortality dimension. Those rules that are determined to be highly weighted are part of the risk assessment rules and are to be measured in every differential monitoring focused review or inspection. There are no exceptions to this.

Key indicator predictor rules is the other methodology which is part of the differential monitoring approach. Key indicator or predictor rules statistically predict overall regulatory compliance and are a very efficient metric for determining the overall regulatory compliance of a facility but in a summary, targeted, and focused fashion without having to do a comprehensive inspection in looking at all the rules and regulations.

Using the combined methodologies of key indicator predictor rules and risk assessment rules makes the differential monitoring approach the most effective and efficient program monitoring system because it focuses on those rules where clients/children may be injured while at the same time predicting overall regulatory compliance with all the rules. It is the perfect balance of effectiveness and efficiency. This is the highly recommended way to utilize differential monitoring, but many jurisdictions/states/provinces use either the risk assessment or the key indicator methodologies, few are utilizing both. Hopefully this will change as the regulatory science field matures over the upcoming decades.

Let's move from the theory, program monitoring approaches and methodologies to the actual measurement of licensing data. Licensing data are at the nominal measurement level. This is

important which will be pointed out shortly in the specific approach being taken here. The approach we will take is to use the Confusion Matrix, which is a well-known metric in the decision-making computer research literature and refocus it for regulatory science within the context of the definition of regulatory compliance and licensing measurement. It will also deal with the policy implications of this particular metric. It is being proposed that this new Uncertainty-Certainty Matrix (UCM) is a fundamental building block to licensing decision making. The 2 x 2 matrix has been written about a great deal in the development of the various methodologies described above and is the center piece for determining key indicator predictor rules, but it is also a core conceptual framework in licensing measurement and ultimately in program monitoring and reviews.

The reason for selecting this matrix is the nature of licensing data, it is binary or nominal in measurement. Either a rule/regulation is in-compliance or out of compliance. Presently most jurisdictions deal with regulatory compliance measurement in this nominal level or binary level. There is to be no gray area, this is a clear distinction in making a licensing decision about regulatory compliance. The UCM also takes the concept of Inter-Rater Reliability (IRR) a step further in introducing an uncertainty dimension that is very important in licensing decision making which is not as critical when calculating IRR. It is moving from an individual metric to a group metric involving regulatory compliance with rules.

The key pieces to the UCM are the following: the decision (D) regarding regulatory compliance and actual state (S) of regulatory compliance. Plus (+) = In-compliance or Minus (-) = Out of compliance. So, let's build the matrix:

Table 1: Uncertainty-Certainty Matrix (UCM) Logic Model

UCM Matrix Logic		**Decision (D) Regarding**	**Regulatory Compliance**
		(+) In Compliance	(-) Not In Compliance
Actual State (S) of	(+) In Compliance	Agreement	Disagreement
Compliance	(-) Not In Compliance	Disagreement	Agreement

The above UCM matrix demonstrates when agreement and disagreement occur which establishes a level of certainty (Agreement Cells) or uncertainty (Disagreement Cells). In a perfect world, there would only be agreements and no disagreements between the decisions made about regulatory compliance and the actual state of regulatory compliance. But from experience, this is not the case based upon reliability testing done in the licensing research field in which a decision is made regarding regulatory compliance with a specific rule or regulation and then that is verified by a second observer who generally is considered the measurement standard.

Disagreements raise concerns in general, but the disagreements are of two types: false positives and false negatives. A false positive is when a decision is made that a rule/regulation is out of compliance when it is in compliance. Not a good thing but its twin disagreement is worse where with false negatives it is decided that a rule/regulation is in compliance when it is out of compliance. False negatives need to be avoided because they place clients at extreme risk, more so than a false positive. False positives should also be avoided but it is more important to deal with the false negatives first before addressing the false positives.

The next logical question after dealing with the measurement issues of licensing data and the fact that it is measured nominally is how best to deal with a data distribution which is severely skewed. In Figure 1, dichotomization was introduced in the graphic in depicting the differences

between high and low regulatory compliance. As presented above in attempting to eliminate false negatives and reduce false positives, the same can be done by dichotomizing the licensing data distribution in order to accentuate the differences between low regulatory compliance and substantial + full regulatory compliance. Dichotomization of data is generally not recommended from a statistical point of view but because of the nature of licensing data being measured at the nominal level and being so severely skewed, it is warranted.

Regulatory Compliance has been always approached as an all or none phenomenon, whether a rule is in-compliance, or it is not. There is no in-between or shades of gray or partial compliance. This worked when the prevailing paradigm was that full regulatory compliance and program quality were a linear relationship. This was the assumption but not empirically verified until the later 1970's-1980's. When this assumption was put to an empirical test, it did not hold up but rather a curvilinear relationship between regulatory compliance and program quality was discovered. This upset the prevailing paradigm and suggested we needed a new approach to addressing the relationship between regulatory compliance and program quality as mentioned earlier in this article.

It became clear after these findings in the 1970's-80's and then in the 2010's when replication studies were completed that substantial regulatory compliance could not be ignored based upon this new theory of regulatory compliance in which substantial compliance acted as a "sweet spot" of best outcomes or results when comparing regulatory compliance and program quality scores. The nominal metric needed to be revised and more of an ordinal metric was to be its replacement. Because now it wasn't just being in or out of compliance, but it mattered which rules were in or out of compliance and how they were distributed. This revised

application involved aggregate rules and does not apply to individual rule scoring. The studies completed between 1970's and 2010's involved aggregate rules and not individual rules. To determine if the nominal to ordinal metric needs to be revised still needs empirical data to back this change.

The introduction of substantial compliance into the regulatory compliance measurement strategy moved the field from an instrument-based program monitoring into a more differential monitoring approach. With differential monitoring this approach considered which rules and how often reviews should be done. Also, a new Regulatory Compliance Scale was proposed to consider the importance of substantial compliance based upon the regulatory compliance theory of diminishing returns. As this Regulatory Compliance Scale has evolved within the licensing health and safety field it needs further revision in which program quality can be infused into the decision making related to individual rules. Remember that the original studies were concerned about rules in the aggregate and not individual rules. It has now become apparent that in dealing with the infusion of quality into rule formulation, a return to the individual rule approach makes the most sense.

The next iteration of the Regulatory Compliance Scale will contain the following categories: Exceeding Full compliance, Full compliance, Substantial compliance, and Mediocre compliance to adjust for the infusion of the quality element. This differs slightly from the original aggregate rule Regulatory Compliance Scale where the categories were Full compliance, Substantial compliance, Mediocre compliance, and Low compliance where only licensing health and safety elements were considered (see the Table 2 below which depicts the regulatory compliance scales and program monitoring systems side by side).

Without the theory of regulatory compliance, differential and integrative monitoring would not be needed because regulatory compliance would have had a linear relationship with program quality and full compliance would have been the ultimate goal. There would have been no need for targeted rule enforcement or reviews because all rules would have had an equal weight when it came to protecting clients and any individual rule would have predicted overall compliance. But it "just ain't so" as it is said. The need to make adjustments is brought about by the theory and it has not been the same ever since.

Table 2: Regulatory Compliance Scales and Program Monitoring Systems

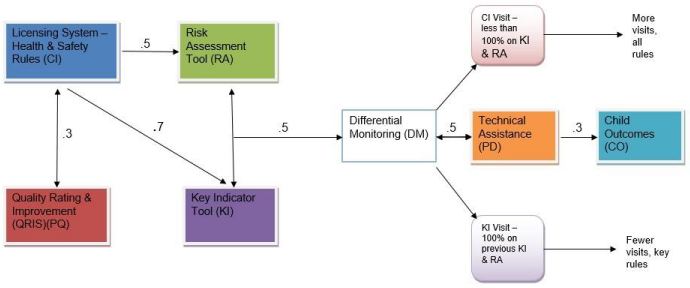
Scoring Level	**Individual Rule**		**Aggregate Rules**	**Individual Rule**
Scale	**Instrument based**	***Scale***	**Differential**	**Integrated**
7	Full Compliance	7	Full Compliance	Exceeds Compliance
-	---	5	Substantial	Full Compliance
-	---	3	Mediocre	Substantial
1	Out of Compliance	1	Low	Mediocre/Low

The above table attempts to summarize in tabular form the previous paragraphs in describing the relationship between program monitoring and licensing measurement scaling via a proposed Regulatory Compliance Scale. As one can see this moves the paradigm from a nominal to an ordinal measurement rubric and depicts the differences in the measurement focus either at the individual rule or aggregate rules scoring levels. It also considers the significance of substantial compliance given the theory of regulatory compliance in which substantial compliance focus is a "*sweet spot*" phenomenon as identified in the regulatory science research literature. It is hoped that the regulatory science field takes these paradigm shifts into consideration in moving forward with building licensing decision making systems and how licenses are issued to facilities.

As a final footnote, keep in mind that the theory of regulatory compliance applies to the relationship between regulatory compliance and program quality and does not apply to regulatory compliance in and of itself related to health and safety. When dealing with regulatory compliance, full compliance is the ultimate goal with individual rules and in determining which rules are predictive rules. It is the preferred methodology in order to eliminate false negatives and decreasing false positives in making licensing decisions related to regulatory compliance.

So, what are the takeaways from the theory of regulatory compliance and its implications for regulatory science.

1) The theory of regulatory compliance has ushered in a new paradigm demonstrating the importance of substantial compliance and putting it on equal footing with full 100% regulatory compliance.
2) Regulatory compliance will not get us to quality on its own, rules and regulations need an infusion of quality so there is the need to balance regulatory compliance and quality standards in any future promulgation of rules and regulations.
3) How does all this fit together? An Early Childhood Program Quality Improvement and Indicator Model has been proposed to build off the results of the theory of regulatory compliance and to build a robust program monitoring system that both differentiates and integrates. See the following Figure 3 which provides a logic model for how the model would play out.



$$\sum CI \times \sum PQ \Rightarrow \sum RA + \sum KI \Rightarrow \sum DM + \sum PD \Rightarrow CO$$

1

Figure 3: Early Childhood Program Quality Improvement and Indicator Model

4) All the studies and research presented in this article are from the human services area. It will be interesting to see if other industries in the medical, scientific, and economic arenas demonstrate the same type of relationship between regulatory compliance in their respective industries and sets of rules and regulations and the ultimate quality of the products they produce.
5) The ceiling effect, diminishing returns, plateauing all depict a curvilinear relationship rather than a linear relationship. As additional studies are completed, this relationship needs to be fine-tuned. Hopefully moving from a nominal measurement strategy to one that is more ordinally based via the Regulatory Compliance Scale will help to fine-tune that relationship.
6) The idiosyncratic nature of licensing data distributions needs to be dealt with statistically because of severe skewness in the data which limits the analytical frames that can be used. Various weighting schemes are being attempted in order to build in

more variance in the data and the infusion of more quality standards into rule formulation should help.

7) Hopefully, this article has given the reader the necessary background to understand this new paradigm for licensing measurement and monitoring systems with all its intricacies and foibles.

References:

Fiene, R. (2023). ***Licensing Measurement and Monitoring Systems: Regulatory Science Applied to Human Service Regulatory Administration***, National Association for Regulatory Administration, Licensing Curriculum, Fredericksburg, Virginia.

Fiene, R. (2022). Regulatory Compliance Monitoring Paradigms and the Relationship of Regulatory Compliance/Licensing with Program Quality: A Policy Commentary. ***Journal of Regulatory Science 10, no. 1, 1-7.*** https://doi.org/10.21423/JRS-V10A239

Fiene, R. (2019). A treatise on theory of regulatory compliance, ***Journal of Regulatory Science, 7, no. 1, 1-3.*** https://doi.org/10.21423/JRS-V07FIENE

Note:

In the regulatory science research literature and in this article both the theory of regulatory compliance as well as the regulatory compliance theory of diminishing returns are used interchangeably.

Richard Fiene, Emeritus Professor of Psychology, Penn State University has had a long career in regulatory science and developmental psychology in which his research has led to the development of national early care and education standards. He directs the research institute, ***Research Institute for Key Indicators Data Lab*** *where he continues to conduct research utilizing his statistical methodologies to determine the key indicator predictors related to early care and education quality. Email: RFiene@RIKInstitute.com*

The Uncertainty-Certainty Matrix Logic Model and Algorithms

Richard Fiene PhD

Penn State Edna Bennett Pierce Prevention Research Center

May 2024

This research abstract will take the Confusion Matrix which is a well-known metric in the decision-making research literature and refocus it on regulatory science within the context of the definition of regulatory compliance and licensing measurement. It will also deal with the policy implications of this particular metric. In this abstract, it is proposed that the Uncertainty-Certainty Matrix (UCM) is a fundamental building block to licensing decision making. The 2 x 2 matrix is the center piece for determining key indicator rules, but it is also a core conceptual framework in licensing measurement and ultimately in program monitoring and reviews.

The reason for selecting this matrix is the nature of licensing data, it is binary or nominal in measurement. Either a rule/regulation is in compliance or out of compliance. Presently most jurisdictions deal with regulatory compliance measurement in this nominal level or binary level. There is to be no gray area, this is a clear distinction in making a licensing decision about regulatory compliance. The UCM also takes the concept of Inter-Rater Reliability (IRR) a step further in introducing an uncertainty dimension that is very important in licensing decision making which is not as critical when calculating IRR. It is moving from an individual metric to a group metric (See Figures 1 & 2) involving regulatory compliance with rules.

The key pieces to the UCM are the following: the decision (D) regarding regulatory compliance and actual state (S) of regulatory compliance. Plus (+) = In-compliance or Minus (-) = Out of compliance. So, let's build the matrix:

Table 1: Uncertainty-Certainty Matrix (UCM) Logic Model

UCM Matrix Logic		**Decision (D) Regarding**	**Regulatory Compliance**
		(+) In Compliance	(-) Not In Compliance
Actual State (S) of	(+) In Compliance	Agreement	Disagreement
Compliance	(-) Not In Compliance	Disagreement	Agreement

The above UCM matrix demonstrates when agreement and disagreement occur which establishes a level of certainty (Agreement Cells) or uncertainty (Disagreement Cells). In a perfect world, there would only be agreements and no disagreements between the decisions made about regulatory compliance and the actual state of regulatory compliance. But from experience, this is not the case based upon reliability testing done in the licensing research field in which a decision is made regarding regulatory compliance with a specific rule or regulation and then that is verified by a second observer who generally is considered the measurement standard.

Disagreements raise concerns in general, but the disagreements are of two types: false positives and false negatives. A false positive is when a decision is made that a rule/regulation is out of compliance when it is in compliance. Not a good thing but its twin disagreement is worse where with false negatives it is decided that a rule/regulation is in compliance when it is out of compliance. False negatives need to be avoided because they

place clients at extreme risk, more so than a false positive. False positives should also be avoided but it is more important to deal with the false negatives first before addressing the false positives.

Let's look at this from a mathematical point of view in the following matrix. In order to better understand the above relationships and determine when ameliorative action needs to occur to shore up the differences between the agreements and disagreements, it is easier to do this mathematically than trying to eyeball it.

Table 2: Uncertainty-Certainty Matrix (UCM) Math Model

UCM Matrix Math Model		**Decision (D) Regarding**	**Regulatory Compliance**	**Totals**
		(+) In Compliance	(-) Not In Compliance	
Actual State (S)	(+) In Compliance	A	B	Y
Of Compliance	(-) Not In Compliance	C	D	Z
Totals		W	X	

Formulae based upon above: Agreements = (A)(D); Disagreements = (B)(C); Randomness = sqrt ((W)(X)(Y)(Z))

UCM Coefficient = ((A)(D)) - ((B)(C)) / sqrt ((W)(X)(Y)(Z)) in which a coefficient closer to 1 indicates agreement (certainty) and a coefficient closer to –1 indicates disagreement (uncertainty). A coefficient closer to 0 indicates randomness. Obviously, we want to see (A)(D) being predominant and very little in (B)(C) which are false positives and negatives where decisions and the actual state of regulatory compliance are not matching. If (WXYZ) is predominant then there is just randomness in the data. Also, not an intended result.

The reason for even suggesting this matrix is the high level of dissatisfaction with the levels of reliability in the results of program monitoring reviews as suggested earlier. If it were not so high, it would not be an issue; but with it being so high the field of licensing needs to take a proactive role in determining the best possible way to deal with increasing inter-rater reliability among licensing inspectors. Hopefully, this organizational schema via the UCM Matrix will help to think through this process related to licensing measurement and monitoring systems.

$$\boldsymbol{UCM} = \ll \boldsymbol{A} \times \boldsymbol{D} \gg - \ll \boldsymbol{B} \times \boldsymbol{C} \gg \ \div \sqrt{\ } \ll \boldsymbol{W} \times \boldsymbol{X} \times \boldsymbol{Y} \times \boldsymbol{Z} \gg$$

The above formula provides a means to calculate when action needs to be taken based upon the respective UCM coefficients. A UCM coefficient from +.25 to +1.00 is in the acceptable range; +.24 to -.24 is due to randomness and needs to be addressed with additional inter-rater reliability training; -.25 to –1.00 indicates a severe disagreement problem that needs to be addressed both in reliability training and a full review of the targeted rules/regulations to determine if the specific rule needs additional clarification.

Table 3: Uncertainty-Certainty Matrix (UCM) Licensing Decision Coefficient Ranges

UCM Coefficient	**Licensing Decision**
+.25 to +1.00	Acceptable, No Action Needed, In or Out of Regulatory Compliance Verified through mostly Agreements. (Generally, 90% of cases)
+.24 to -.24	Random, Agreements + Disagreements, Needs Reliability Training. (Generally, 5% of cases)
-.25 to –1.00	Unacceptable, Mostly Disagreements, Needs Training & Rule/Regulation Revision. (Generally, 5% of cases)

Figure 1: Kappa Coefficient

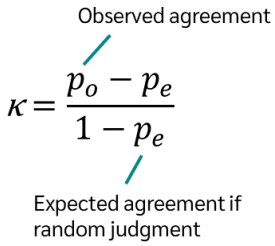


Figure 2: Uncertainty-Certainty Coefficient

$$\phi = \frac{ad - bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

$$\phi = \sqrt{\frac{\chi^2}{n}}$$

Let's provide an example of how this could work. A standard/rule/regulation that is common is the following:

Do all caregivers/teachers and children wash their hands often, especially before eating and after using the bathroom or changing diapers?

This is obviously an observation item where the licensing staff would observe in a sample of classrooms in a child care center for a set period of time. During their observations, there were several opportunities where the necessary behavior was required, and the staff complied with the rule and washed their hands. So, on the surface this specific rule was in compliance and there would appear to be full compliance with this rule based upon the observation.

A second scenario is where the observation is made, and the licensing staff observes the child care staff not washing their hands on several occasions. Then this specific rule would be out of compliance, and it would be duly noted by the licensing staff. These two scenarios establish a certain level of certainty during this observation session. However, there are other outcomes, for example, possibly one of the classrooms that was not observed had the opposite finding than what was observed in these particular classrooms. If data were being aggregated and a specific percentage was to be used the final decision about this rule could be different. Now we are getting into the uncertainty cells of the matrix where a false positive or negative could be the result. The licensing staff records the rule as being in compliance when in reality it is not = false negative or the rule is recorded as being out of compliance when in reality it is in compliance = false positive.

Another example which involves either Random Clinical Trials (RCT) or the use of abbreviated inspections (AI) and the results from these two interventions. The decision making in both RCT and AI is

basically the same. We want to make sure that the results match reality. Every time an abbreviated review is done the following four regulatory compliance results should occur based upon the UCM matrix: 1) no additional random non-compliance is found; 2) there are no false negatives (abbreviated review finds no non-compliance but in reality there is); 3) when there is non-compliance found in abbreviated inspections, other related non-compliance is found; and 4) lastly the level of false positives (abbreviated review finds non-compliance but in reality there are no other related non-compliances) is kept to a minimum. This last result based upon copious research is that it is difficult to obtain but as the regulatory science moves forward hopefully this will become more manageable.

Hopefully these above examples provided some context for how the Uncertainty-Certainty Matrix (UCM) can be used in making specific licensing decisions based upon the regulatory compliance results.

1. Uncertainty-Certainty Matrix for Validation and Reliability Studies

The purpose of this part of this research abstract is to explore the possibility of utilizing the Uncertainty-Certainty Matrix (UCM) in validation and reliability studies in licensing decision making. The UCM has been proposed for use in licensing decision making but this would be an extension of this thinking to studies that involve validating licensing decisions such as when key indicators are used in comparison with comprehensive reviews of rules, and in reliability studies to determine individual inspector bias in regulatory compliance.

The basic premise of the UCM is that individual decision-making matches reality. When it comes to regulatory compliance decision making a 2 x 2 matrix can be drawn with the possible outcomes as is indicated in the following table (Table 4).

Table 4

UCM Matrix Logic		**Decision Regarding**	**Regulatory Compliance**
		(+) In Compliance	(-) Not In Compliance
Actual State of	(+) In Compliance	Agreement (++)	Disagreement (+-)
Compliance	(-) Not In Compliance	Disagreement (-+)	Agreement (--)

In using this table, the hope is that the decision regarding regulatory compliance matches the actual state of compliance where the coefficient is as close to +1.0 as possible, in other words, perfect agreement. So, the agreement cells are heavily weighted. We do not want to see all the cells, both agreement and disagreement cells, equally weighted. That would indicate a random response rate and a coefficient close to 0.0.

But there is another possibility which involves bias on the part of the licensing inspector in which they have certain biases or tendencies when it comes to making regulatory compliance decisions about individual rules. So, it is possible that decisions made regarding regulatory compliance could be either overall (+) positive In-Compliance or (-) negative Not-In-Compliance when in reality the actual state of compliance is more random.

When this occurs, the coefficient falls off the range category and is not between 0 and +/-1.0 because there is no variance detected in the data. It is always biased either positively or negatively.

The UCM can be used for both reliability and validity testing as suggested in the above. Just look for different results. For validity, false positives and negatives should either be eliminated or reduced as well as possible and the remaining results should show the typical diagonal pattern as indicated by the agreement cells.

For reliability, the same pattern should be observed as in the validity testing above but there is an additional test in which bias is tested for. Bias will be ascertained if the patterns in the results indicate a horizontal or vertical pattern in the data with little or no diagonal indication. Bias can be found at the individual inspector level as well as at the standard level or the actual state of compliance.

In both reliability and validity testing, random results in which each of the cells are equally filled is not a desirable result either.

The following tables 5-10 depict the above relationships with results highlighted in red:

Table 5

Valid & Reliable Results	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Table 6

Random Results	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Table 7

Positive Bias Results Individual	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Table 8

Negative Bias Results Individual	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Table 9

Positive Bias Results Standard	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Table 10

Negative Bias Results Standard	(+) In Compliance	(-) Not In Compliance
(+) In Compliance	Agreement (++)	Disagreement (+-)
(-) Not In Compliance	Disagreement (-+)	Agreement (--)

Tables 5 – 10 demonstrate the different results based upon individual response rates when making regulatory compliance decisions about rules. Table 5 is what needs to be attained and tables 6 – 10 need to be avoided. Only in table 5 are false negatives and positives eliminated or avoided. In tables 6 – 10, false negatives and/or false positives are introduced which is not desirable when making validity or reliability decisions.

Table 6 results clearly indicate that a great deal of randomness has been introduced in the regulatory compliance decision making in which the individual licensing inspector decisions do not match reality. Tables 7 and 8, demonstrate bias in the decision-making process either positively (inspector always indicates in compliance) or negatively (inspector always indicates out of compliance). It is also possible that the standard being used has bias built into it, this is less likely but is still a possibility. The results in Tables 9 and 10 demonstrate where this could happen.

All these scenarios need to be avoided and should be monitored by agency staff to determine if there are patterns in how facilities are being monitored.

2. Uncertainty-Certainty Matrix for Differential Monitoring Studies

The purpose of this part of the research abstract is to explore the possibility of utilizing the Uncertainty-Certainty Matrix (UCM) not only in validation and reliability studies in licensing decision making but also with differential monitoring studies. The UCM has been proposed for use in licensing decision making but this would be an extension of this thinking to studies that involve validating licensing decisions such as when key indicators are used in comparison with comprehensive reviews of rules, and in the development of risk rules as part of the risk assessment methodology. This new Differential Monitoring 2x2 Matrix can also be used to depict the relationship between full and substantial regulatory compliance and the nature of rulemaking.

The basic premise of the DMM: Differential Monitoring Matrix is similar to the original thinking with the UCM but there are some changes in the formatting of the various cells in the matrix (see Table 11). When it comes to regulatory compliance decision making a 2 x 2 matrix can be drawn with the possible outcomes as is indicated in Table 11 where each individual rule is either in (+) or out (-) of compliance. Also, there is the introduction of a high regulatory compliant group (+) and a low regulatory compliant group (-) which is different from the original UCM.

Table 11

DMM Matrix	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

By utilizing the format of Table 11, several key components of differential monitoring can be highlighted, such as key indicators and risk assessment rules, as well as the relationship between full and substantial regulatory compliance.

Regulatory compliance is grouped into a high group (+), generally this means that there is either full or substantial regulatory compliance with all rules. The low group (-) usually has 10 or more regulatory compliance violations. Individual rules being in (+) or out (-) of regulatory compliance is self-explanatory.

Tables 12-18 below will demonstrate the following relationships:

Table 12 depicts the key indicator relationship between individual rules and the high/low groups as indicated in red. In this table, the individual rule is in compliance with the high group and is out of compliance with the low group. This result occurs on a very general basis and should have a .50 coefficient or higher with a p value of less than .0001.

Table 13 depicts what most rules look like in the 2x2 DMM. Most rules are always in full compliance since they are standards for basic health and safety for individuals. This is especially the case with rules that have been weighted as high-risk rules. Generally, one never sees non-compliance with these rules. There will be a substantial number of false positives (+-) found with high-risk rules but that is a good thing.

Table 14 depicts what happens when full compliance is used as the only criterion for the high group. Notice that the cell right below (++) is eliminated (-+). This is highly recommended since it eliminates false negatives (-+) from occurring in the high group. As will be seen in Table 5, when substantial compliance is used as part of the high group sorting, false negatives are re-introduced. If possible, this should be avoided, however in some cases because of the regulatory compliance data distribution it is not always possible where not enough full compliant programs are present.

Table 15 depicts what occurs when substantial compliance is used as part of determining the high group. False negatives can be reintroduced into the matrix which needs to be either eliminated or reduced as best as possible. If substantial compliance needs to be used in determining the high group, then there is a mathematical adjustment that can be made which will impact the equation and essentially eliminate false negatives mathematically (see the research note at the end of this research abstract).

Table 16 depicts what happens if the individual rule is particularly difficult to comply with. Both the high performers as well as the low performers are out of compliance with the rule.

Table 17 depicts a situation where the programs are predominantly in a low group with few at full or substantial regulatory compliance which is indicative of poor performing programs. Very honestly, this is generally not seen in the research literature, but it is a possibility and one to be in tune with.

Table 18 depicts a terrible individual rule which predicts just the opposite of what we are trying to do with programs. Obviously, this rule would need to be rewritten so that it fits with the essence of regulatory compliance in helping to protect individuals.

The following tables 12-18 will depict the above relationships with results highlighted in red:

Table 12

Key Indicators	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Table 13

Risk Rules	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Table 14

Full Compliance	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance		(--)

Table 15

Substantial Compliance	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Table 16

Very Difficult Rule	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Table 17

Poor Performing Programs	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Table 18

Terrible Rule	High Group (+)	Low Group (-)
(+) Rule is In Compliance	(++)	(+-)
(-) Rule is Not In Compliance	(-+)	(--)

Tables 12 – 18 demonstrate the different results based on the relationship between individual regulatory compliance and if a program is either a high performer or a low performer. These tables are provided as guidance for understanding the essence of differential monitoring and regulatory compliance which has various nuances when it comes to data distributions. This research abstract hopefully can be used as a guide in determining from a data utilization point of view how to make important regulatory compliance policy decisions, such as: which rules

are excellent key indicator rules, which are performing as high risk rules, importance of full compliance, what to do when substantial compliance needs to be employed, are there difficult rules to comply with, how well are our programs performing, and do we have less than optimal rules that are in need of revision.

Research Note:

Over the past decade in doing research on the Regulatory Compliance Key Indicator Metric (RCKIm) it has become very clear that false negatives needed to be controlled for because of their potential to increase morbidity and mortality. When dealing with regulatory compliance and full compliance as the threshold for the high grouping variable in the 2 x 2 Regulatory Compliance Key Indicator Matrix (RCKIm) (see matrix below in Table 19), false negatives could be either eliminated or reduced to the point of no concern.

However, if substantial compliance rather than full compliance is used as the threshold for the high grouping variable in the 2 x 2 Regulatory Compliance Key Indicator Matrix (RCKIm) this becomes a problem again. There is the need to introduce a weighting factor. In utilizing the RCKIm, the following equation/algorithm is used to produce the Fiene Coefficient (FC):

$$FC = ((A)(D)) – ((B)(C)) / sqrt (WXYZ)$$

This RCKIm needs to be revised/updated to the following to consider the need to again eliminate false negatives being generated by the results of the equation/algorithm; this can be accomplished by cubing B:

$$FC^* = ((A)(D)) – ((B^3)(C)) / sqrt (WXYZ)$$

By this simple adjustment to cube (B = False Negatives) it will basically eliminate the use of any results in which a false negative occurs when substantial compliance is determined. The table below (Table 19) displays the variables of the Regulatory Compliance Key Indicator Matrix (RCKIm).

Table 19: RCKIm	High RC Group	RC Low Group	
KI In Compliance	A	B^3	Y
KI Violations	C	D	Z
Totals	W	X	

In the above examples, FC can be used when the High RC Group is at full regulatory compliance, but FC* needs to be used when the High RC Group is including substantial as well as full regulatory compliance. By using both equations/algorithms, it better deals with the results of the Regulatory Compliance Theory of Diminishing Returns.

The results should clearly show that only positive (+) coefficients will become Regulatory Compliance Key Indicators versus those rules that do not show any relationship to overall regulatory compliance (0), but now the negative (-) coefficients will more clearly show when any false negatives appear and clearly not include them as Regulatory Compliance Key Indicators. This is a major improvement in the Regulatory Compliance Key Indicator methodology which clearly demonstrates the differences in the results. It provides a gateway in regulatory compliance data distributions where substantial regulatory compliance is heavily present while full regulatory compliance is not. This could become a problem as the regulatory science field moves forward with the use of the Regulatory Compliance Theory of Diminishing Returns.

3. Uncertainty-Certainty Matrix Logic Model and Algorithms

The Uncertainty-Certainty Matrix (UCM) has been introduced for licensing decision making and validating differential monitoring approaches. This section builds upon this introduction by presenting a logic model and its corresponding algorithms. The purpose of this section is to expand that original thinking and to introduce the associated algorithms that specify the UCMLMA (Uncertainty-Certainty Matrix Logic Model and Algorithms).

The UCMLMA will build off the contingency table and confusion matrix logic modeling as described above (in tables 1 and 2) in the following matrix (Table 20).

Table 20: UCMLMA: Uncertainty-Certainty Matrix Logic Model and Algorithms

UCMLMA	Actual:	Individual	Inspector
Reality:	Decisions:	C = Compliance	NC = Non Compliance
Gold	C = Compliance	*TP = True Positive*	*FP = False Positive*
Standard	NC = Non Compliance	*FN = False Negative*	*TN = True Negative*

From this matrix certain licensing decision making can be made and certain biases can be avoided with the following algorithms related to compliance and non compliance of rules/regulations.

TP + TN is the ideal situation in which all decisions are made correctly. Accuracy can be measured by *((TP x TN) - (FN x FP) / sqrt ((TP + FN) (FP + TN) (TP + FP) (FN + TN))* with a coefficient closer to +1.00.

TP + FP + FN + TN is a totally random situation where decisions are not reliable nor valid, there is no rhyme or reason to the decision making process by the individual inspector. Randomness can be measured by the following: *(TP x TN) - (FN x FP) / sqrt ((TP + FN) (FP + TN) (TP + FP) (FN + TN))* with a coefficient closer to 0.00.

TP + FP introduces a positive bias in which the inspector has a tendency to always make a decision in which the program is in compliance with specific rules. They are overly lenient in their interpretation of the rules. Positive bias sensitivity can be measured by *FP / (TP + FP).* The higher the percent, the more bias is present.

FN + TN introduces a negative bias in which the inspector has a tendency to always make a decision in which the program is in non compliance with specific rules. They are overly stringent in their interpretation of the rules. Negative bias sensitivity can be measured by *FN / (FN + TN).* The higher the percent, the more bias is present.

Table 21: UCMLMA: Uncertainty-Certainty Matrix Logic Model and Algorithms Applied to Differential Monitoring

UCMLMA		Overall	Compliance
Individual	Decisions:	High Group (top 10%)	Low Group (bottom 10%)
Rule/Standard/	C = Compliance	*TP = True Positive*	*FP = False Positive*
Regulation	NC = Non Compliance	*FN = False Negative*	*TN = True Negative*

From this matrix (Table 21), certain decisions can be made regarding differential monitoring validation related to key indicators, risk assessment, regulatory compliance levels and individual rule performance with the following algorithms.

TP + TN is the ideal result when determining key indicators because the individual rule statistically predicts overall compliance with all rules. Accuracy can be measured with the following: *((TP x TN) - (FN x FP) / sqrt ((TP + FN) (FP + TN) (TP + FP) (FN + TN))* with the coefficient being closer to +1.00.

TP + FP is what happens with high risk rules in that they are always in compliance and one sees very little non compliance with these high risk rules generally. Sensitivity can be measured by *FP / (TP + FP)*. Higher compliance rates would generally indicate higher risk rules.

TP is when 100% full compliance is always present. The greater the number or percent of total programs being 100% in full compliance, the greater the skewness in the data distribution.

TP + FN is present when substantial regulatory compliance is used as a criterion for licensing decision making. The higher the number, the greater the skewness in the data distribution.

FN + TN occurs when a rule is difficult to comply with. The following algorithm can be used *FN / (FN + TN)* to measure how difficult the rule is. The greater the percent, the more difficult the rule is.

FP + TN occurs when programs are performing poorly and there is a great deal of non compliance with the rules. The greater the number of programs indicates a poorly performing system of rules.

FP + FN is an example of a terrible rule that predicts the opposite of what we intended with overall regulatory compliance. The following algorithm *((TP x TN) - (FN x FP) / sqrt ((TP + FN) (FP + TN) (TP + FP) (FN + TN))* can be used with coefficients closer to -1.00 demonstrating this finding.

The above two tables (table 20 and 21) and their corresponding algorithms give a structured approach to licensing decision making and differential monitoring validation. It should provide the licensing administrator with a data driven and empirical method based upon regulatory science principles. By utilizing the concepts, logic model, and algorithms presented in this working paper, it will hopefully add to a regulatory scientific approach in the human services licensing and regulatory administration field when it comes to making licensing decisions that avoid bias and reliability & validity errors in decision making.

Potential Solution to the Child Care Trilemma Revisited: It is not about how many rules or how few rules, it is about finding the "Right Rules", The Holy Grail of Early Care and Education

Richard Fiene PhD

Penn State Edna Bennett Pierce Prevention Research Center

August 2024

Recently in the early care and education field, there has been a great deal of discussion about deregulation of early care and education standards/rules/regulations in order to have increased access to child care which has become even more of a problem since the COVID19 Pandemic. This discussion or controversy has been going on for a long time, it is nothing new. Either we have too few rules/regulations/standards or we have too many. The pendulum has swung both ways depending on the political will and grass roots advocacy efforts. But I think there is a better way to deal with this discussion which is driven by the regulatory science and the empirical evidence that has emerged over the past 50 years. Let's take this discussion out of the political domain and place it where it needs to be, firmly within the newly emerging regulatory science field and focus on regulatory compliance.

So where do we start? In a very familiar place, with *Caring for Our Children* which is the default set of health and safety standards in the early care and education field. Over the past 30+ years it has been the reference for child care licensing agencies as they think about promulgating new or revised rules/regulations/standards in their respective states. It is based upon the latest science in developmental psychology, pediatrics, and public health fields related to early care and education settings. It has gone through 4 editions so we know it is up to date. But it is a daunting document, 700+ standards are within this reference manual for the early childhood field.

Advocates point to *Caring for Our Children* as the go-to document because it provides a solid floor to quality while building on this base to demonstrate best practices. Others, mostly in the political arena, point to it as an example of over-regulation, too many rules to follow. But let's not forget what *Caring for Our Children* is all about, protecting our children while in out of home care. Yes, child care is critical to parents who need to work, but it is equally important (many will argue that it is more important) that our children are safe while in out of home care. This is a delicate balance and all part of the child care trilemma of availability, affordability and quality. As we adjust one, the others are impacted proportionately or disproportionately depending on your vantage point.

So, what is a potential solution to the child care trilemma? Let's look at regulatory science for potential guidance. As I said earlier, regulatory science is an emerging field, it is not well developed as the other physical and social sciences but it is making tremendous strides in the past 20-30 years. Albeit, mostly in the pharmaceutical industry but some very interesting innovations have occurred in the human services, in particular in early care and education. As the title suggests, a course of action should be not over-regulating and not deregulating but identifying the "right rules/regulations/standards". Within the regulatory science field this has occurred in early care and education and it applies directly to *Caring for Our Children*. It was evident to the original designers of *Caring for Our Children* that there was a problem with the sheer volume of the set of standards. The designers started to look at ways to distill it down to a more manageable size. And this is where regulatory science came into play with a new methodology that was emerging around risk assessment and key indicator rules/regulations.

Initially there was more focus on the risk assessment methodology to determine if certain *Caring for Our Children* standards placed children at increased risk of morbidity and mortality if regulatory non-compliance occurred. The resulting document*, Stepping Stones to Caring for Our Children*, came about based upon this risk assessment rule methodology. It took the 700+ standards to distill it down to approximately 120 standards. It became a much more manageable document that state licensing agencies could use.

Later in the development and evolution of *Stepping Stones to Caring for Our Children*, again borrowing from the regulatory science field, the key indicator rule methodology was utilized to determine if there were a smaller set of standards that had more of a predictive value in protecting children when it came to regulatory compliance in an overall sense. This resulted *in Caring for Our Children Basics* (approximately 65 standards) which was originally proposed as a voluntary set of standards for all early care and education. I think it was a good idea back when it was first proposed and I still think it is a good idea.

In fact, let's begin there in attempting to address a revised solution to the child care trilemma. In this discussion about where the child care field is headed and the most recent call for deregulation, let's pivot and think about using *Caring for Our Children Basics* as our point of discussion rather than arbitrarily removing rules with this deregulation mind set because it is politically expedient. Let's be driven by the empirical evidence and the science which *Caring for Our Children Basics* is derived from, solid regulatory compliance methodologies of risk assessment and key indicator rule/regulatory/standard identification. See how your state's child care rules size up with *Caring for Our Children Basics* in making sure that at the very least all these standards are in place. Templates from regulatory science have been developed to do this comparison.

Then once this is done in the aggregate, begin to look at the individual standards within *Caring for Our Children Basics*. Let's be honest, probably the most discussed standard is Staff-Child Ratios and Group Sizes. It has the greatest impact on cost (staff), numbers (children), and quality. This has been clearly demonstrated in the research literature over the past 50+ years. Nothing has changed, it was the focal point back in the 1970's and it is today. But let's think outside the regulatory compliance box for a minute and maybe we don't look at staff-child ratios in isolation but cross it with another standard/rule such as the qualifications of staff and suggest an alternate rule where staff-child ratio can be increased slightly but only with the most highly qualified staff?! Like I said, let's think outside the regulatory compliance box. And while we're there, the fee that is attained by the program with the additional child should go to the more qualified staff as an add on to their salary. Yes, they have an additional child but they also have the revenue generated as a salary increase with the addition.

As with staff-child ratio and group size, we perform the same type of critical analysis utilizing the empirical regulatory compliance data available to make changes in the existing set of rules. As has been pointed out in the regulatory science research literature, regulatory compliance with rules is a measurement issue, so it should be solved in a corresponding way, use the data, don't ignore it and leave it up to the whims of the political process to determine what stays and what gets pitched.

The point of this research abstract position paper is for us to take a step back and avoid a knee-jerk reaction to dealing with the child care crisis and that the only solution is to increase availability and affordability at the expense of health and safety via deregulation. We now have an emerging regulatory science to guide us and I hope we use it for making educated and informed choices as we move forward in attempting to solve the continuing child care trilemma.

Richard Fiene PhD, Research Psychologist/Regulatory Scientist, Edna Bennett Pierce Prevention Research Center, Penn State University,
rfiene@rikinstitute.com

Please do not quote, submitted for publication in ***Exchange*** magazine.

Development of a Regulatory Compliance Scale

Richard Fiene PhD

Penn State's Edna Bennett Pierce Prevention Research Center

April 2025

The purpose of this paper is to provide an alternate paradigm for regulatory compliance measurement in moving from a nominal to an ordinal scale measurement strategy and to introduce a new licensing/regulatory compliance metric: the Regulatory Compliance Scale. Regulatory compliance measurement is dominated by a nominal scale measurement system in which rules are either in compliance or out of compliance. There are no gradients for measurement within the present licensing measurement paradigm. It is very absolute. Either a rule is in full compliance to the letter of the law or the essence of the regulation or it is not. An alternate paradigm borrowing from accreditation and other program quality systems is to establish an ordinal scale measurement system which takes various gradients of compliance into account. With this alternate paradigm, it offers an opportunity to begin to introduce a quality element into the measurement schema. It also allows us to take into consideration both risk and prevalence data which are important in rank ordering specific rules.

So how would this look from a licensing decision making vantage point. Presently, in licensing measurement, licensing decisions are made at the rule level in which each rule is either in or out of compliance in the prevailing paradigm. Licensing summaries with corrective actions are generated from the regulatory compliance review. It is a nominal measurement system being based upon Yes/No responses. The alternate measurement paradigm I am suggesting in this paper is one that is more ordinal in nature where we expand the Yes/No response to include gradients of the particular rule. In the next paragraph, I provide an example of a rule that could be measured in moving from a nominal to ordinal scale measurement schema.

Rather than only measuring a rule in an all or none fashion, this alternate paradigm provides a more relative mode of measurement at an ordinal level. For example, with a professional development or training rule in a particular state which requires, let's say, 6 hours of training for each staff person. Rather than having this only be 6 hours in compliance and anything less than this is out of compliance, let's have this rule be on a relative gradient in which any amount of hours above the 6 hours falls into a program quality level and anything less than the 6 hours falls out of compliance but at a more severe level depending on how far below the 6 hours and how many staff do not meet the requirement (prevalence). Also throw in a specific weight which adds in a risk factor, and we have a paradigm that is more relative rather than absolute in nature.

From a math modeling perspective, the 1 or 0 format for a Yes or No response becomes -2, -1, 0, +1, +2 format. This is more similar to what is used in accreditation systems where 0 equals Compliance and -1 and -2 equals various levels of Non-Compliance in terms of severity and/or prevalence. The +1 and +2 levels equal value added to the Compliance level by introducing a Quality Indicator. This new formatting builds upon the compliance vs non-compliance dichotomy (C/NC) but now adds a quality indicator (QI) element. By adding this quality element, we may be able to eliminate or at least lessen the non-linear relationship between regulatory compliance with rules and program quality scores as measured by the Environmental Rating Scales (ERS) and CLASS which is the essence of the Theory of Regulatory Compliance (TRC). It could potentially make this a more linear relationship by not having the data as skewed as it has been in the past.

By employing this alternate paradigm, it is a first demonstration of the use of the Key Indicator Methodology in both licensing and quality domains. The Key Indicator Methodology has been utilized a great deal in licensing but in few instances in the program quality domain. For example, over the past five years, I have worked with approximately 10 states in designing Licensing Key Indicators but only one state with Quality Key Indicators from their QRIS – Quality Rating and Improvement System. This new paradigm would combine the use in both. It also takes advantage of the full ECPQI2M – Early Childhood Program Quality Improvement and Indicator Model by blending regulatory compliance with program quality standards.

A major implication in moving from a nominal to an ordinal regulatory compliance measurement system is that it presents the possibility of combining licensing and quality rating and improvement systems into one system via the Key Indicator Methodology. By having licensing indicators and now quality indicators that could both be measured by licensing inspectors, there would be no need to have two separate systems but rather one that applies to everyone and becomes mandated rather than voluntary. It could help to balance both effectiveness and efficiency by only including those standards and rules that statistically predict regulatory compliance and quality and balancing risk assessment by adding high risk rules.

I will continue to develop this scale measurement paradigm shift in future papers but wanted to get this idea out to the regulatory administration field for consideration and debate. This will be a very controversial proposal since state regulatory agencies have spent a great deal of resources on developing free standing QRIS which build upon licensing systems. This alternate paradigm builds off the Theory of Regulatory Compliance's key element of relative vs absolute measurement and linear vs non-linear relationships (Fiene, 2022). Look for additional information about this on RIKI Institute Blog - https://rikinstitute.com/blog/.

Introduction to the Regulatory Compliance Scale

The theory of regulatory compliance has been proven in multiple studies over the past four decades and has been utilized extensively in the creation of differential monitoring and its spin off methodologies of risk assessment and key indicators (Fiene, 2025). In fact, differential monitoring would not have been possible without the theory of regulatory compliance because

the paradigm which it replaced, one of one-size-fits-all monitoring or uniform monitoring would have predominated. However, with the theory of regulatory compliance which introduced the importance of substantial regulatory compliance and the search for the right rules/regulations that made a difference in client's lives, rather than emphasizing more or less regulations or rules.

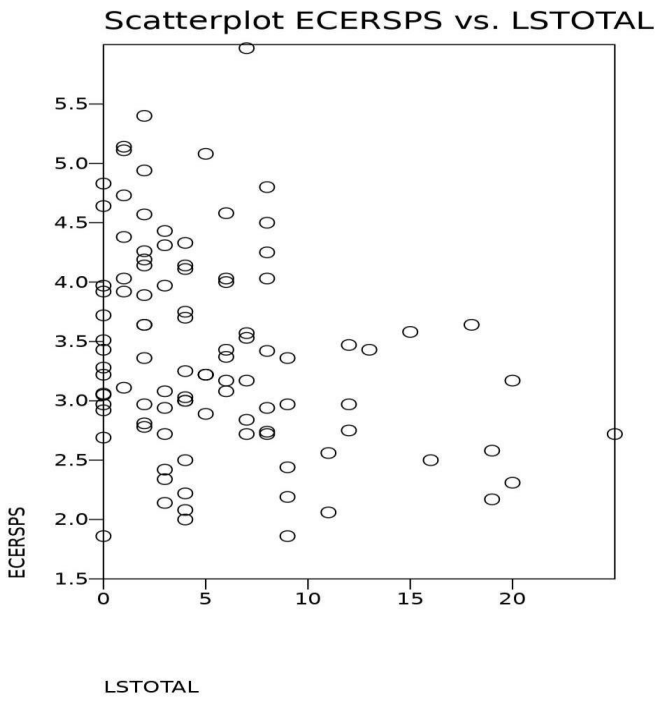
The theory of regulatory compliance has another application when it comes to regulatory compliance measurement in helping to move the licensing field from a nominal based measurement strategy to one of ordinal based measurement. The new measurement strategy is the Regulatory Compliance Scale (RCS) and it is depicted in the following table.

RCS	***Compliance***	***Risk***	***Model***	***Model***
Scale	***Level***	***Level***	***Violations***	***Weights***
7 = A	**Full**	**None**	**0**	**0**
5 = B	**Substantial**	**Low**	**1-3**	**1-3**
3 = C	**Medium**	**Medium**	**4-9**	**4-6**
1 = D	**Low**	**High**	**10+**	**7+**

The above table needs some explanation. The first column is the proposed ordinal scale similar to other scales utilized in the program quality measurement research literature on a 1 – 7 Likert Scale where 7 = Full Regulatory Compliance, 5 = Substantial Regulatory Compliance, 3 = Medium Regulatory Compliance, and 1 = Low Regulatory Compliance. It could also be thought of as an Alpha Scale of A – D as well. The next column has the compliance levels that run from full 100% regulatory compliance to low regulatory compliance. The third column depicts the risk level from none to high which corresponds with the compliance levels. The next two columns depict two models, one unweighted and one in which the rules are weighted with corresponding weights. These models are based upon the two prevailing approaches to rank ordering rules or regulations in the research literature.

The following figures will depict how the scale was conceived based upon empirical evidence in the various studies supporting the theory of regulatory compliance.

The first figure shows the actual individual violation data of the programs compared to their corresponding ECERS scores. There is not a significant relationship between the two as depicted in the graphic.

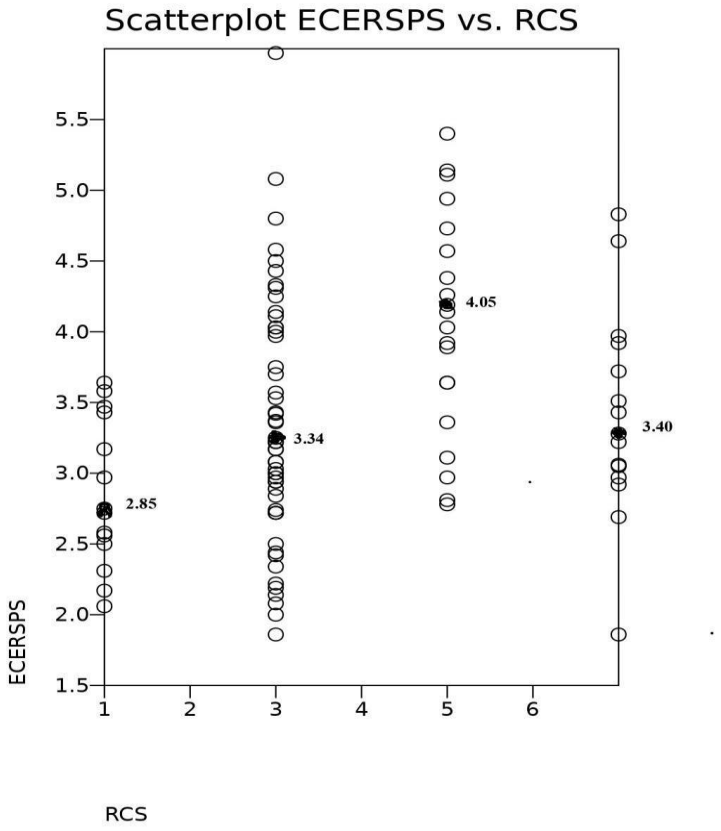


The following figure below depicts what occurs when the individual violation data are grouped according to the theory of regulatory compliance in which a substantial compliance category is introduced, and the data are moved from a nominally based metric to an ordinally based metric of full, substantial, medium, and low regulatory compliance categories. This grouping more clearly reflects the theory of regulatory compliance. It also clearly demonstrates the ceiling effect which is an outcome of the theory of regulatory compliance in which substantial and full regulatory compliance levels are basically equivalent when quality is taken into account. Or at the extreme level which is depicted here where full regulatory compliance quality scores are actually lower than the substantial regulatory compliance quality scores. A footnote about the figures and the scaling: the scales for the first figure are on a lower to higher progression but the higher LSTOTAL represents higher non-compliance where the second figure is also based upon lower to higher but the higher scores represent increased quality and increased regulatory compliance.

So, in reading the change from left to right, these two figures are reversed images of each other. This is just a quirk of the scaling and not a mistake in the plotting of data.

The RCS has been pilot tested in both the non-weighted and weighted models and based upon these studies it appears to be more effective in distinguishing quality amongst the various categories rather than utilizing violation count data. This would be a significant improvement when it comes to licensing measurement. Of course, additional replication studies need to be

completed before it would be recommended as a new Scale to be used for making licensing decisions.



The above figure is dramatically different than the prevailing paradigm which predicts a linear relationship between regulatory compliance and quality which is the paradigm of a uniform monitoring approach. The above results clearly indicate a reconsideration with the introduction of substantial regulatory compliance as an important contributor to overall quality if not the most important contributor to quality. As stated above, these findings have been replicated in several studies conducted over the past several decades.

This would be a major paradigm shift in moving from individual violation data counts to an ordinal scale metric but it does warrant additional research. The problem with individual violation data is that it doesn't take into account the relative risk of the individual rule which could place clients at increased risk of morbidity or mortality. Risk assessment has worked really well when coupled with key indicators in the differential monitoring approach and it appears to be an asset in the development of a Regulatory Compliance Scale (RCS).

Regulatory Compliance Scale Studies

The Regulatory Compliance Scale (RCS) was introduced several years ago and has been used in a couple of validation studies for differential monitoring and regulatory compliance's ceiling effect phenomenon. RCS buckets or thresholds were statistically generated based upon these studies, but it is time to validate those buckets and thresholds to determine if they are really the best model in creating a regulatory compliance scale. Since proposing the RCS, there has been a great deal of interest from jurisdictions in particular from Asian and African nations. Additional statistically based trials were conducted, and this brief report is the compilation of those trials over the past year.

The data used are from several jurisdictions that are part of the international database maintained at the Research Institute for Key Indicators Data Laboratory at Penn State University focusing on program quality scores and rule violation frequency data. These data from the respective databases were recoded into various thresholds to determine the best model. The jurisdictions were all licensing agencies in the US and Canada geographically dispersed where both regulatory compliance and program quality data was obtained from a sample of early care and education programs.

Methodology

The following methodology was used starting with the original RCS buckets/thresholds of Full, Substantial, Medium, and Low regulatory compliance:

RCS Models used for analyses

RCS				**Models**			
		Original	***1***	***2***	***3***	***4***	***5***
	Full	100	100	100	100	100	100
Scaling	***Substantial***	99-98	99-97	99-97	99-98	99-98	99-97
	Medium	97-90	96-90	96-93	97-95	97-85	96-85
	Low	89>	89>	92>	94>	84>	84>

Five alternate models were used to compare the results to the original RCS. The numbers indicate the number of violations subtract from a perfect score of 100. Full regulatory compliance indicates no violations and a score of 100 on the scale. The next bucket of 99-98 indicates that there were 1 or 2 regulatory compliance violations which resulted in a 99-98 score on the scale. This logic continues with each of the models.

The scale score was determined in the following manner: Full Regulatory Compliance = 7; Substantial Regulatory Compliance = 5; Medium Regulatory Compliance = 3; and Low Regulatory Compliance =1. This rubric is how the original RCS scaling was done on a Likert type scale similar to other ECE program quality scales, such as the Environmental Rating Scales.

Results

The following results are correlations amongst the respective RCS Models from Table above compared to the respective jurisdictions program quality tool (Quality1-3): ERS or CLASS Tools.

RCS Model Results compared to Quality Scales

RCS results	Models	Quality1	Quality2	Quality3
Jurisdiction1	**RCS0**	**.26***	**.39***	**.39***
	RCS3	.21	.32*	.33*
	RCS5	.20	.36*	.33*
Jurisdiction2	**RCS0**	**.76****	**.46****	**---**
	RCS3	.12	-.07	---
	RCS5	.18	-.02	---
	RCSF1	.55**	.29*	---
	RCSF2	.63**	.34	---
Jurisdiction3	RCS0	.19	.18	.16
	RCS3	**.21**	**.21**	**.15**
	RCS5	.18	.16	.07
	RCSF1	.17	.17	.10
	RCSF2	.18	.18	.19
Jurisdiction4	RCS0	.24*	---	---
	RCS3	.28*	---	---
	RCS5	**.30***	**---**	**---**
	RCSF1	.21	---	---
	RCSF2	.29*	---	---
Jurisdiction5	RCS0	.06	-.02	.07
	RCS3	.06	-.01	.05
	RCS5	**.08**	**.00**	**.09**
	RCSF1	.00	-.03	.05
	RCSF2	.05	-.03	.05

*Statistically significant .05 level;

**Statistically significant .01 level.

In the above table starting under Jurisdiction2, two new models were introduced based upon the Fibonacci Sequence (Fibonacci1 = RCSF1; Fibonacci2 = RCSF2) and their model structure is in the following Table. The reason for doing this is that the Fibonacci Sequence introduces additional variation into the scaling process.

RCS Fibonacci Models

RCS Fibonacci			**Models**	
		Original	***Fibonacci1***	***Fibonacci2***
	Full	100	100	100
Scaling	***Substantial***	99-98	40	90
	Medium	97-90	20	20
	Low	89>	13	13

A second series of analyses were completed in comparing the RCS models with program quality (Quality1) by running ANOVAs with the RCS models as the independent variable and program quality as the dependent variable. The reason for doing this was the nature of the data distribution in which there was a ceiling effect phenomenon identified which would have had an impact on the correlations in table above. All results are significant at $p < .05$ level with the exception of Jurisdiction2.

ANOVAs Comparing the RCS Models with Program Quality

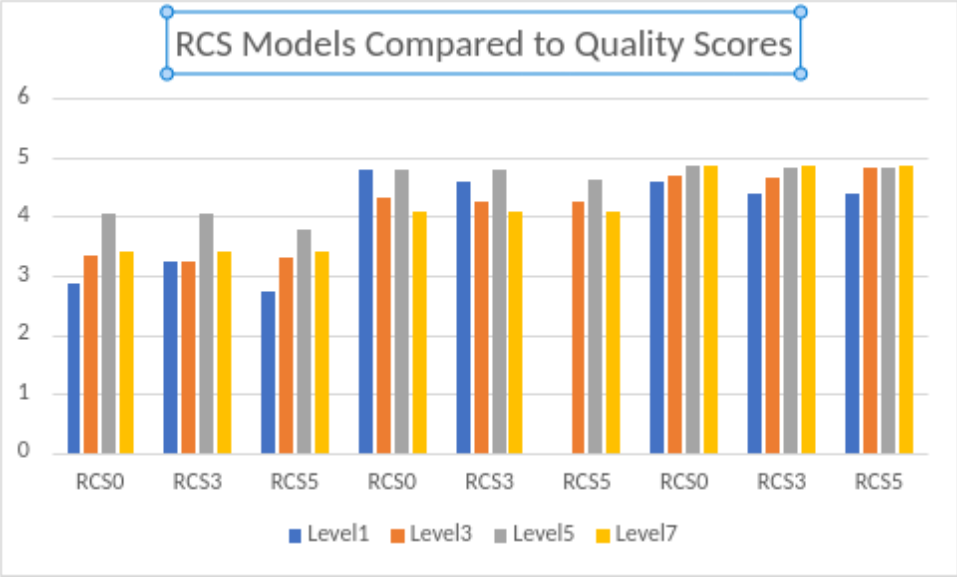
Jurisdictions	**Model**	**Level 1**	**Level 3**	**Level 5**	**Level 7**
Jurisdiction1	**RCS0**	**2.85**	**3.34**	**4.05**	**3.40**
	RCS3	3.24	3.23	4.05	3.40
	RCS5	2.73	3.32	3.77	3.40
Jurisdiction2	RCS0	4.81	4.31	4.80	4.10
	RCS3	**4.59**	**4.25**	**4.80**	**4.10**
	RCS5	---	4.26	4.64	4.10
Jurisdiction3	RCS0	4.59	4.68	4.86	4.87
	RCS3	**4.38**	**4.67**	**4.83**	**4.87**
	RCS5	4.38	4.83	4.83	4.87
Jurisdiction4	RCS0	37.81	37.01	44.28	41.96
	RCS3	36.57	38.60	44.28	41.96
	RCS5	**33.46**	**36.53**	**43.10**	**41.96**
Jurisdiction5	RCS0	3.93	4.17	4.28	4.07
	RCS3	4.02	4.24	4.28	4.07
	RCS5	**3.75**	**4.13**	**4.26**	**4.07**

Discussion

Based upon the above results, it appears that the original RCS model proposed in 2021 is still the best model to be used, although the Fibonacci Sequence model is a close second in some of the jurisdictions. This model will need further exploration in determining its efficacy as a replacement or enhancement to the original RCS Model.

The bottom line is that the original RCS Model is as good as any and no other model is consistently better than all the rest. The RCS Model does have a slight edge over Regulatory Compliance Violation RCV frequency counts in some jurisdictions but not in others. It is much easier to interpret the relationship between quality and the RCS models than it is to interpret the results from the quality scores and the RCV data distribution. So, the recommendation would be for licensing agencies to think about using this new scaling technique in one of its model formats to determine its efficacy. Pairing up RCS and RCV data side by side by licensing agencies would be important studies to determine which approach is the better approach.

The below graphic depicts the relationship between the RCS Models (0, 3, 5) when compared to the quality scores (1-6) clearly showing the ceiling effect and diminishing returns effect phenomenon so typical of regulatory compliance data when compared to program quality. These graphs are from the first three jurisdictions (1, 2, 3) from the above tables.



Additional Analyses Comparing the 11 Studies

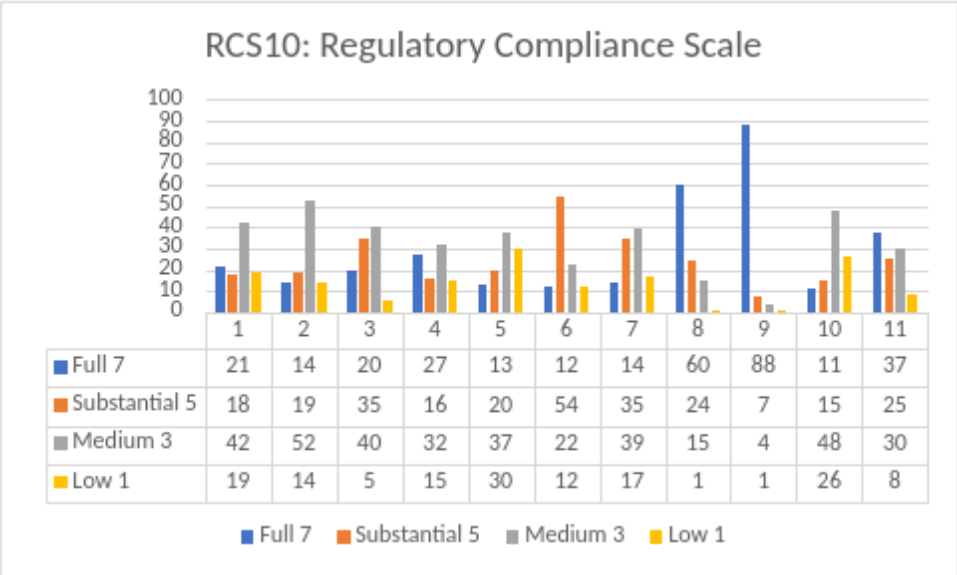
This section provides the results from 11 studies from 10 states and Canadian Provinces in which the proposed new Regulatory Compliance Scale (RCS) was utilized as a byproduct of a differential monitoring implementation or validation study. These studies were undertaken over a decade long period (2013-2023).

The RCS was based upon the following rubric: Full Regulatory Compliance (100%) or no violations = 7; Substantial Regulatory Compliance (99-98) or 1-2 violations = 5; Medium Regulatory Compliance (97-90) or 3-10 violations =3; and Low Regulatory Compliance (89 or less) or 11 or more violations = 1.

These are the results from these 10 jurisdictions which are presented in the following Table (all results are presented as percents of programs that fell into the scaling 1-7). Under the Studies, the number of the specific study is provided, followed by the sample size, followed by if it is in the USA (US) or Canada (CA).

RCS Scale			**RCS Scaling**		
Studies	**7=Full**	**5=Substantial**	**3=Medium**	**1=Low**	**Comments**
1-403-US	21%	18%	*42%*	19%	***High Med NC***
2-104-US	14%	19%	*52%*	14%	***High Med NC***
3-422-US	***20%***	***35%***	*40%*	***5%***	***OK***
4-219-CA	***27%***	***16%***	*32%*	***15%***	***OK***
5-60-CA	13%	20%	*37%*	***30%***	***High NC/Low C***
6-585-US	***12%***	***54%***	*22%*	***12%***	***OK***
7-255-US	***14%***	***35%***	*39%*	***17%***	***OK***
8-1399-US	***60%***	24%	15%	***1%***	***Low NC/High C***
9-2116-US	***88%***	7%	4%	***1%***	***Low NC/High C***
10-482-US	11%	15%	*48%*	*26%*	***High NC/Low C***
11-3070-US	***37%***	***25%***	*30%*	***8%***	***OK***

In looking at the results, it is preferrable to have most of the programs at either a full or substantial regulatory compliance level (7 or 5) and to have fewer programs at the medium or low regulatory compliance level (3 or 1). But in those jurisdictions where there are higher percentages of programs at the medium or low levels of regulatory compliance, it could be that their enforcement of rules and regulations is more stringent. This potential result needs further investigation to get to the root cause of these differences because there is a good deal of variation across the jurisdictions as is evident from the above table.



	1	2	3	4	5	6	7	8	9	10	11
Full 7	21	14	20	27	13	12	14	60	88	11	37
Substantial 5	18	19	35	16	20	54	35	24	7	15	25
Medium 3	42	52	40	32	37	22	39	15	4	48	30
Low 1	19	14	5	15	30	12	17	1	1	26	8

Based upon the above studies and results, the regulatory compliance scale (Fiene, 2022) which appears from recent studies to be a better metric in measuring regulatory compliance than just counting the number of violations that a program has related to their respective rules, regulations, or standards. So how does the regulatory compliance scale work. It essentially puts violations into buckets of regulatory compliance as follows: full compliance (100%) or no violations; substantial compliance (99-98%) or 1-2 violations; mediocre compliance (97-90%) or 3-9 violations; and lastly low/non-optimal compliance (89% or lower) or 10+ violations. Why buckets, because logically it works, it is the way we think about regulatory compliance. It is a discrete rather than continuous metric and logically fits into these four categories. This is based upon 50 years of research into regulatory compliance data distributions and when the data are moved from frequency counts of violation data into these buckets/categories, the math works very well in identifying the better performing programs.

Regulatory Compliance Scale Extensions

Depicted below is a regulatory compliance grid model showing the relationship between regulatory compliance (RC) and program quality (PQ).

An explanation of the below chart will demonstrate how regulatory compliance and program quality in human service facilities interact. The horizontal blue axis depicts the various levels of regulatory compliance while the vertical green axis depicts the various levels of program quality of facilities. It ranges from 1-5 or low to high for each axis. The red "**X's**" represents the relationship that has been identified in the research literature based upon the theory of regulatory compliance in which there is either a plateau effect or a downturn in quality as regulatory compliance increases. The one italicized "***X***" is an outlier that has also been identified in the research literature in which some (it does not happen often) low compliant programs really are at a high-quality level.

It is proposed in order to mitigate the plateau effect with regulatory compliance and program quality standards because regulatory compliance data distributions are severely skewed which means that many programs that have questionable quality are being included in the full (100%) compliance domain. When regulatory compliance standards are increased in their quality components this will lead to a higher level of overall quality as depicted in the "**XX**" cell all the way on the lower right. It also helps to mi gate the severe skewness in the regulatory compliance data distribution. The data distribution does not approximate a normally distributed curve which is the case with the program quality data distribution.

Regulatory Compliance x Program Quality Grid Model

PQ/RC ->	1 Low	2 Med	3 Substantial	4 Full 100%	5QualityAdditions
1 Low	XXX				
2		XX			
3 Med			XX	XXX	
4			XX	X	
5 High	*X*				XX

By utilizing this model, it helps to deal more directly in taking a non-linear relationship and making it linear again when comparing regulatory compliance with program quality. This model provides a theoretical approach supporting what many state licensing administrators are thinking from a policy standpoint: add more quality to health and safety rules/regulations. This grid/matrix also depicts the three regulatory compliance models: Linear, Non-linear, and Stepped.

Here is another potential extension of the Regulatory Compliance Scale using the ECPQIM DB – Early Childhood Program Quality Improvement and Indicator Model Data Base, it is possible to propose developing and using a Regulatory Compliance Scoring System and Scale (RC3S). This new proposed RC3S could be used by state human service agencies to grade facilities as is done in the restaurant arena. Presently, in the human service field, licenses are issued with a Certificate of Compliance but generally it does not indicate what the regulatory compliance level is at. This new proposal would alleviate this problem by providing a scale for depicting the level of regulatory compliance.

The ECPQIM DB is an international data base consisting of a myriad group of data sets drawn from around the USA and Canada. It has been in the making over 40 years as of this writing, so its stability and generalizability have been demonstrated. What follows is the chart depicting the RC3S.

Regulatory Compliance Scoring System and Scale (RC3S)

Color	Non-Compliance Level	Regulatory Compliance Level
Blue	**0**	**Full Compliance**
Green	**1-2**	**Substantial Compliance**
Yellow	**3-6**	**Mid-Range Compliance**
Orange	**7-9**	**Low Compliance**
Red	**10-15+**	**Very Low Compliance**

It is evident from the above chart that the color go from blue to red which indicates an increasing risk of non-compliance and a lower level of overall regulatory compliance, which is not a good thing in the licensing field. Non-compliance levels indicate the number of rules or regulations or standards that are not complied with. And lastly, the regulatory compliance level indicates the movement from full (100% regulatory compliance with all rules) to very low compliance with rules. These ranges for the scaling are based on 40 years of research in understanding and plotting the data distributions around the world related to regulatory compliance in the human services. These results have consistently appeared over this 4-decade time period and show no signs of changing at this point.

Regulatory Compliance Scaling for Decision Making

There is a lack of empirical demonstrations of regulatory compliance decision making. In the past, I have used the methodologies of key indicators, risk assessment and the resultant differential monitoring techniques of how often and what should be reviewed for decision making. What has not been addressed is decision making based upon comprehensive reviews when all regulations are assessed. This section addresses how empirical evidence taken from the past 40+ years of establishing and researching a national database for regulatory compliance can help lead us to a new scaling of regulatory compliance decision making.

In analyzing regulatory compliance data, it becomes perfectly clear that the data have very little variance and are terribly skewed in which the majority of programs are in either full or substantial compliance with all the respective regulations. Only a small handful of programs fall into the category of being in low compliance with all the regulations.

The proposed scaling has three major decision points attached to regulatory compliance scores. Either programs are in full or substantial compliance, in low compliance or somewhere in the middle. Full or substantial regulatory compliance is 100% or 99-98% in regulatory compliance. Low regulatory compliance is less than 90% and mid-regulatory compliance is between 97%-90%. These ranges may seem exceptionally ght but based upon the national database on regulatory compliance that I maintain at the Research Institute for Key Indicators (RIKILLC) these are the ranges that have formed over the past 40 years. These data ranges should not come as a surprise because we are talking about regulatory compliance with health and safety standards. These are not quality standards; these are basic protections for clients. The data are not normally distributed, not even close as is found in quality tools and standards.

What would a **Regulatory Compliance Decision-Making Scale** look like:

Data	Level	Decision
100-98%	***Full/Substantial***	***License***
97-90%	***Mid-Range***	***Provisional***
89% or less	***Low***	***No-License***

States/Provinces/Jurisdictions may want to adjust these levels, and the scaling based upon their actual data distribution. For example, I have found certain jurisdictions to have very unusually skewed data distributions which means that these ranges need to be ghten even more. If the data distribution is not as skewed as the above scale, then these ranges may need to be more forgiving.

This regulatory compliance decision making scale does not take into account if abbreviated methodologies are used, such as risk assessment or key indicator models that are used in a differential monitoring approach. The above scale is to be used if a jurisdiction decides not to use a differential monitoring approach and wants to measure regulatory compliance with all regulations and complete comprehensive reviews.

Conclusion

The Theory of Regulatory Compliance (Fiene, 2019) and bringing substantial compliance to the fore front of regulatory science has been written about a great deal. This paper builds upon these previous assertions and expands them into some practical applications that can be utilized within regulatory science as it relates to licensing measurement, regulatory compliance scaling, and monitoring systems paradigms. This paper has introduced the Regulatory Compliance Scale which is a departure in how best to measure regulatory compliance. This new scale along with the proposed Uncertainty-Certainty Matrix (Fiene, 2025b) provides a robust licensing measurement and program monitoring strategy. This paper provides the last piece of a differential monitoring approach that includes instrument-based program monitoring, key indicators, risk assessment, and the uncertainty-certainty matrix.

Regulatory Compliance has been always approached as an all or none phenomenon, whether a rule is in compliance, or it is not. There is no in-between or shades of gray or partial compliance. This worked when the prevailing paradigm was that full regulatory compliance and program quality were a linear relationship. This was the assumption but not empirically verified until the later 1970's-1980's. When this assumption was put to an empirical test, it did not hold up but rather a curvilinear relationship between regulatory compliance and program quality was discovered. This upset the prevailing paradigm and suggested we needed a new approach to addressing the relationship between regulatory compliance and program quality.

It became clear after these findings in the 1970's-80's and then in the 2010's when replication studies were completed that substantial regulatory compliance could not be ignored based upon this new theory of regulatory compliance in which substantial compliance acted as a "sweet spot" of best outcomes or results when comparing regulatory compliance and program quality scores. The nominal metric needed to be revised and more of an ordinal metric was to be its

replacement. Because now it wasn't just being in or out of compliance, but it mattered which rules were in or out of compliance and how they were distributed. This revised application involved aggregate rules and does not apply to individual rule scoring. The studies completed between 1970 and 2010 involved aggregate rules and not individual rules. To determine if the nominal to ordinal metric needs to be revised still needs empirical data to back this change.

The introduction of substantial compliance into the regulatory compliance measurement strategy moved the field from an instrument-based program monitoring into a more differential monitoring approach. With differential monitoring this approach considered which rules and how often reviews should be done. Also, a new Regulatory Compliance Scale was proposed to take into account the importance of substantial compliance based upon the regulatory compliance theory of diminishing returns. As this Regulatory Compliance Scale has evolved within the licensing health and safety field it needs further revision in which program quality can be infused into the decision making related to individual rules. Remember that the original studies were concerned about rules in the aggregate and not individual rules. It has now become apparent that in dealing with the infusion of quality into rule formulation, a return to the individual rule approach makes the most sense.

The next iteration of the Regulatory Compliance Scale will contain the following categories: Exceeding full compliance, Full compliance, Substantial compliance, and Mediocre compliance to adjust for the infusion of the quality element. This differs slightly from the original aggregate rule Regulatory Compliance Scale where the categories were Full compliance, Substantial compliance, Mediocre compliance and Low compliance where only licensing health and safety elements were considered (see the Table below which depicts the regulatory compliance scales and program monitoring systems side by side).

Without the Theory of Regulatory Compliance, differential and integrative monitoring would not be needed because regulatory compliance would have had a linear relationship with program quality and full compliance would have been the ultimate goal. There would have been no need for targeted rule enforcement or reviews because all rules would have had an equal weight when it came to protecting clients and any individual rule would have predicted overall compliance. But it "just ain't so" as it is said. The need to make adjustments is brought about by the theory and it has not been the same ever since.

Regulatory Compliance Scales and Program Monitoring Systems

Scoring Level	**Individual Rule**		**Aggregate Rules**	**Individual Rule**
Scale	**Instrument based**	***Scale***	**Differential**	**Integrated**
7	Full Compliance	7	Full Compliance	Exceeds Compliance
-	---	5	Substantial	Full Compliance
-	---	3	Mediocre	Substantial
1	Out of Compliance	1	Low	Mediocre/Low

The above table attempts to summarize in tabular form the previous paragraphs in describing the relationship between program monitoring and licensing measurement scaling via a proposed regulatory compliance scale. As one can see this moves the paradigm from a nominal to an ordinal measurement rubric and depicts the differences in the measurement focus either at the

individual rule or aggregate rules scoring levels. It also considers the significance of substantial compliance given the theory of regulatory compliance in which substantial compliance focus is a "*sweet spot*" phenomenon as identified in the regulatory science research literature. It is hoped that the regulatory science field takes these paradigm shifts into consideration in moving forward with building licensing decision making systems and how licenses are issued to facilities.

As a final footnote, keep in mind that the Theory of Regulatory Compliance applies to the relationship between regulatory compliance and program quality and does not apply to regulatory compliance in and of itself related to health and safety. When dealing with regulatory compliance, full compliance is the ultimate goal with individual rules and in determining which rules are predictive rules. It is the preferred methodology in order to eliminate false negatives and decreasing false positives in making licensing decisions related to regulatory compliance.

These above concepts all relate to the field of regulatory compliance and how to make informed decisions about licensing, particularly in the context of program monitoring. Here's how they connect:

Regulatory Compliance Scales:

These scales move away from a binary "compliant" or "non-compliant" approach to regulations. Instead, they acknowledge degrees of compliance, recognizing that minor deviations may not be as detrimental as major ones.

They provide a framework for evaluating the severity and frequency of non-compliance, allowing for more nuanced licensing decisions.

Instrument Based Program Monitoring (IBPM):

This is the traditional method of monitoring compliance, relying on standardized instruments and checklists to assess adherence to specific rules.

It's a comprehensive approach, but can be time-consuming and inflexible, potentially leading to over-regulation or missing important aspects of program quality.

Differential Monitoring (DM):

This approach takes into account the risk associated with different regulations, focusing monitoring efforts on areas with the highest potential for harm or non-compliance.

It allows for a more efficient use of resources and can be tailored to the specific needs of each program.

DM often utilizes Regulatory Compliance Scales to determine the severity of non-compliance and guide the level of monitoring needed.

Integrative Monitoring Systems (IMS):

These systems go beyond simply checking compliance and aim to assess the overall quality of a program.

They integrate data from various sources, including IBPM, DM, and other program-specific metrics, to provide a holistic picture of performance.

IMS can inform licensing decisions by considering not only compliance but also program effectiveness in achieving its goals.

Here's a simplified analogy to illustrate the relationships:

Think of regulations as traffic rules.

IBPM is like a police officer checking every car for every violation, regardless of severity.

DM is like a police officer focusing on patrolling areas with high accident rates or known reckless drivers.

Regulatory Compliance Scales are like different levels of fines based on the severity of the traffic violation.

IMS is like a traffic management system that collects data on accidents, traffic flow, and road conditions to optimize traffic flow and safety.

Relationships:

RCS forms the foundation for DM and IMS by providing a way to assess degrees of compliance.

IBPM provides data for RCS and can be incorporated (with adaptations) into DM and IMS.

DM builds on RCS and IBPM by differentiating the intensity of monitoring based on risk and compliance.

IMS is the most comprehensive approach, integrating RCS, IBPM, DM, and additional data sources for a deeper understanding of program performance.

Regulatory Compliance Scales can be used within any of the monitoring approaches to provide a more nuanced assessment of compliance.

IBPM can be a starting point for differential monitoring, providing data on rule compliance to inform risk assessments.

Differential monitoring can be integrated into an integrative monitoring system, along with other data sources, to provide a comprehensive picture of program performance.

Here are some additional points to consider:

The choice of the most appropriate approach will depend on the specific context, such as the type of program being regulated and the available resources.

Implementation of these alternative paradigms requires careful planning and training of regulators and program providers.

Ongoing research and evaluation are needed to refine these approaches and ensure their effectiveness.

These alternative paradigms offer a more flexible and effective approach to licensing decisionmaking compared to the traditional IBPM approach. They allow for a better

understanding of program strengths and weaknesses, optimize resource allocation, and ultimately lead to better regulatory outcomes.

These concepts offer a shift from traditional "one-size-fits-all" compliance models to more flexible and nuanced approaches that consider risk, program quality, and degrees of compliance. This can lead to more efficient and effective regulatory systems that support program improvement while protecting public safety.

Ultimately, these concepts offer alternative paradigms for licensing decision-making, moving away from a rigid "one-size-fits-all" approach to a more nuanced and risk-based system that considers both compliance and program quality.

References:

Fiene, R. (2019). A treatise on Regulatory Compliance. ***Journal of Regulatory Science, Volume 7***, 2019. https://doi.org/10.21423/jrs-v07fiene

Fiene (2022a). Regulatory Compliance Monitoring Paradigms and the Relationship of Regulatory Compliance/Licensing with Program Quality: A Policy Commentary. (2022). ***Journal of Regulatory Science, 10*(1). https://doi.org/10.21423/JRS-V10A239**

Fiene (2022b). Regulatory Compliance Scale, ***RIKINotes Blog***, January 2022.

Fiene (2023a). ***Licensing Measurement & Monitoring Systems***, Research Institute for Key Indicators, Elizabethtown, Pennsylvania.

Fiene (2023b). Ceiling Effect/Diminishing Returns, Regulatory Compliance Scale, and Quality Indicators Scale, ***Mendeley Data,*** doi: 10.17632/gc423hprcs.1

Fiene (2025a). Finding the Right Rules. ***American Scientist, Volume 113, 1***. pps 16-19.

Fiene (2025b). The Uncertainty-Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies, ***Knowledge Journal***, under review.

NARA (2023). ***Saskatchewan Differential Monitoring/Quality Indicators Scale Validation Study***, National Association for Regulatory Administration, Fredericksburg, Virginia.

Richard Fiene PhD, Senior Research Psychologist, Research Institute for Key Indicators (RIKILLC); Professor of Psychology (ret), Penn State University; Senior Research Consultant, National Association for Regulatory Administration (NARA). Email: Rfiene@RIKInstitute.com Website: http://RIKInstitute.com

Submitted for Publication: ***International Critical Childhood Policy Studies Journal***

The Child Care and Early Education Heart Monitor: The Intersection of Structural Quality and Process Quality Using the Contact Hour Metric As The Foundation

Richard Fiene PhD 0000-0001-6095-5085

Research Institute for Key Indicators Data Laboratory

Penn State Edna Bennett Pierce Prevention Research Center

July 2025

Abstract

The Child Care and Early Education (CCEE) Heart Monitor (CCEEHM) is introduced as a new Integrated Program Monitoring System's Approach to assessing both structural and process quality in one platform. It builds upon the Contact Hour (CH) metric and the Key Indicator Methodology (KIM) that have been introduced in the CCEE licensing and monitoring field. The CCEEHM expands the use of the CH and KIM methods by integrating key elements from both structural and process quality into a software application that can be used by staff, licensors, and quality assessors. The CCEEHM draws indicators from licensing, regulatory compliance, quality rating and improvement systems, and other quality initiatives, such as accreditation, and professional development and technical assistance systems.

The Child Care and Early Education (CCEE) field needs a means to monitor the key elements of structural and process quality in a unified means. The theory of regulatory compliance has been suggested as a unifying framework for structural and process quality (Fiene, 2019; 2021; 2025a,b); but at a more practical level what could be used to essentially unify the monitoring and measurement of both structural and process quality. Generally, structural and process quality are measured separately from each other by using very separate and distinct tools

utilized by licensing inspectors and quality observers (Kontos & Fiene, 1987). This research paper will build off several concepts that deal with the creation of a new Contact Hour (CH) metric replacing measuring compliance with adult-child ratios while unifying structural quality with process quality. With this new unification of structural and process quality, it will help to build a more Integrated Monitoring Systems Approach (Freer & Fiene, 2023) which should go a long way in complementing the measurement strategies employed in licensing and quality rating and improvement systems that have proliferated in the child care and early education field.

Let's begin by placing some context on the title of this new Child Care and Early Education Heart Monitor. What do we mean by heart monitor? Within the research literature in determining the levels of quality generally these levels are broken into two distinctive categories, those that deal with structural quality, such as staff child ratios, group size, etc. Essentially health and safety or licensing rules and regulations. The interactions amongst the staff and children generally fall under the process quality side of the equation. But this is really the "heart" of quality. This is where the magic occurs, the so-called "dance" between the adult and the child(ren). All the structural quality rules and regulations are important in protecting children and keeping them healthy but the interaction of child and adult is where the action occurs. So what is being proposed is to combine these two categories of quality together into one system, placing the measurement and the monitoring of process quality squarely within the structural measurement strategy, the Contact Hour (CH) metric. This will be developed within this paper by fully describing the Contact Hour metric and a newly created CCEE Quality Indicator tool that will measure the quality enhancements within the Contact Hour metric and do this within an

App (software application) that can be downloaded and it will produce the scores based upon reviewing specific documents and observations within a child care and early education program. This new Child Care and Early Education Heart Monitor (CCEEHM) should be both cost effective and efficient being based upon the key indicator methodology (Fiene & Nixon, 1985) and having it developed into an App (software application) should make it particularly easy to use for licensors, assessors, or observers since all the scoring would be done by the CCEEHM App.

Let's continue by delving into the Contact Hour (CH) metric (Fiene & Stevens, 2021). The Contact Hour metric has been proposed as a more effective and efficient metric for measuring compliance with adult-child ratios and group sizes in CCEE programs. It is simple to apply by just asking 6 questions about when children arrive and leave a CCEE program and how many staff are present in a particular classroom (See the second section for the questions and algorithms). Once that is done a trapezoidal model is built in which compliance with staff child and group size rules can be determined. Regulatory compliance is determined by comparing the resultant area to an ideal level of contact between staff and children. This introductory section is followed by the tool that would be used for determining the Contact Hour metric (Section 2) as well as the Program Quality Indicators (PQI)(Section 3) that need to be measured. Also, there is a screen shot of the opening page of the CCEEHM App that has been designed to measure compliance with the tools for CH and PQI at the end of Section 3.

In determining the results, the Contact Hours (CH) are dealt with as absolute values but let's enhance this result by moving it from an absolute value to one that is more relative by introducing process quality measures such as the Program Quality Indicators (PQI). The PQI portion of the tool has a good deal of observations that need to be made in classrooms. To do

this, it would take 1000's of observations to fill the Contact Hour trapezoidal model which is not realistic. But let's let Artificial Intelligence (AI) do the observing and training of AI in what constitutes the various quality levels on the respective CH/PQI tool. By using AI and having video cameras in each of the classrooms to be assessed, this becomes doable. The CH/PQI observer would be able to collect the data by observing and assessing what it sees via the video cameras installed in the classrooms. Summary measurements would be made on an hourly basis and recorded as part of the Contact Hour trapezoidal model. At the end of the day, there would be a relative value utilized in this model rather than the absolute value that has been used in the past to determine structural quality compliance with adult-child ratio and group size. For example, if a CCEE program classroom exceeded the area of the trapezoidal model it would be out of compliance and if it were within the area of the trapezoidal model it was in compliance (see Section 2). By adding the PQI data, it changes this metric totally by adding process quality measures which can be measured on a 1-4 ordinal scale, similar to accreditation systems or an ordinal (1-7) scale, similar to many program quality tools, such as the Environmental Rating Scales (see Section 3).

This approach will get at the ***Heart of CCEE monitoring, "process quality"***, measuring the interactions amongst staff and children in an ongoing fashion. It moves the needle from being structural to process quality providing an intersection of both components of quality. The AI approach will also help to address the issues related to bias in regulatory compliance observing and decision making by inspectors/observers. By training the AI PQI Observers there should be greater certainty established in making the right decisions related to specific quality elements

(Fiene, 2025c). Just as in establishing inter-rater reliability with human observers, the same can be done with the PQI AI Observers but there will be less drift with AI.

The next section describes the Contact Hour Metric methodology in detail. Section 3 provides the Program Quality Indicators (PQI) that are part of the CCEEHM App. These two sections are the meat of the new Integrated Program Monitoring Systems Approach. In fact a human observer could use these two sections and then manually use the CCEEHM App for doing their data entry. The App would then do all the scoring for the individual assessor (See Section 3).

Section 2: Contact Hour (CH) Metric

One starts the Contact Hour (CH) metric methodology by asking the following six questions (The six questions should be asked of each grouping that is defined by a classroom or a well-defined group within each classroom tied to a specific adult-child ratio.):

1.When does your first teaching staff arrive or when does your facility open (TO1)?
2.When does your last teaching staff leave or when does your facility close (TO2)?
3.Number of teaching/caregiving staff (TA)?
4.Number of children on your maximum enrollment day (NC)?
5.When does your last child arrive (TH1)?
6.When does your first child leave (TH2)?

After getting the answers to these questions, the following formulae can be used to determine contact hours (CH) based upon the relationship between when the children arrive and leave (TH) and how long the facility is open (TO):

$$CH = ((NC\ (TO + TH)) / 2) / TA;$$

$$CH = (NC \times TO) / TA;$$

$$CH = ((NC \times TO) / 2) / TA;$$

$$CH = (NC^2) / TA$$

Where: CH = Contact Hours; NC = Number of Children; TO = Total number of hours the facility is open (TO2 - TO1); TA = Total number of teaching staff, and TH = Total number of hours at full enrollment (TH2 - TH1).

By knowing the number of contact hours (CH) it will be possible to rank order the exposure time of adults with children. Theoretically, this metric could then be used to determine that the greater contact hours is correlated with the increased non-regulatory compliance with adult-child ratios as determined in the below table (Table 1).

Table 1: Contact Hour (CH) Conversion Table (RS Model(1.0)) (Fiene, 2020©)

Taking into Account Exposure Time and Density

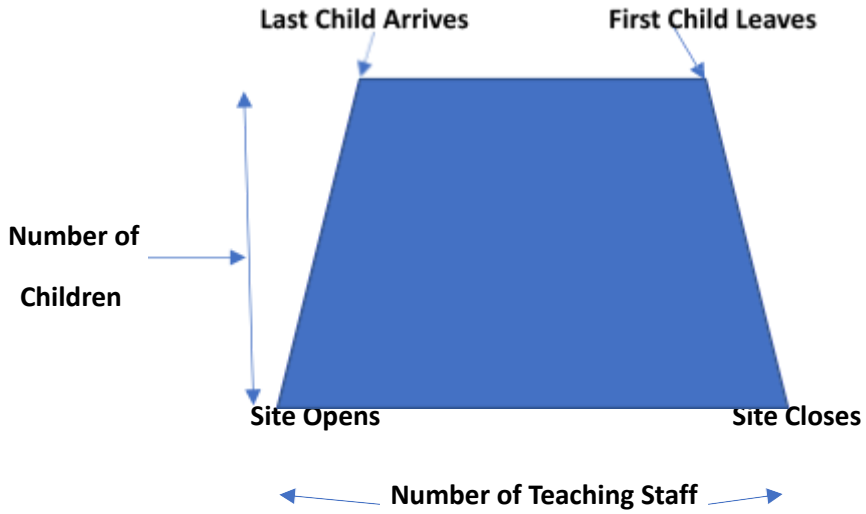
Group Size, Staff Child Ratio, Number of Children and Staff

<---------- Adult-Child Ratios (Relatively Weighted Contact Hours)---------->

NC	CH	1:1	2:1	3:1	4:1	5:1	6:1	7:1	8:1	9:1	10:1	11:1	12:1	13:1	14:1	15:1
1	8	**8**	8	8	8	8	8	8	8	8	8	8	8	8	8	8
2	16	**8**	**16**	16	16	16	16	16	16	16	16	16	16	16	16	16
3	24	**8**	12	**24**	24	24	24	24	24	24	24	24	24	24	24	24
4	32	**8**	**16**	16	**32**	32	32	32	32	32	32	32	32	32	32	32
5	40	**8**	13	20	20	**40**	40	40	40	40	40	40	40	40	40	40
6	48	**8**	**16**	**24**	24	24	**48**	48	48	48	48	48	48	48	48	48
7	56	**8**	14	19	28	28	28	**56**	56	56	56	56	56	56	56	56
8	64	**8**	**16**	21	**32**	32	32	32	**64**	64	64	64	64	64	64	64
9	72	**8**	14	**24**	24	36	36	36	36	**72**	72	72	72	72	72	72
10	80	**8**	**16**	20	27	**40**	40	40	40	40	**80**	80	80	80	80	80
11	88	**8**	15	22	29	29	44	44	44	44	44	**88**	88	88	88	88
12	96	**8**	**16**	**24**	**32**	32	**48**	48	48	48	48	48	**96**	96	96	96
13	104	**8**	15	21	26	35	35	52	52	52	52	52	52	**104**	104	104
14	112	**8**	**16**	22	28	37	37	**56**	56	56	56	56	56	56	**112**	112
15	120	**8**	15	**24**	30	**40**	40	40	60	60	60	60	60	60	60	**120**
16	128	**8**	**16**	21	**32**	32	43	43	**64**	64	64	64	64	64	64	64
17	136	**8**	15	23	27	34	45	45	45	68	68	68	68	68	68	68
18	144	**8**	**16**	**24**	29	36	**48**	48	48	**72**	72	72	72	72	72	72
19	152	**8**	15	22	30	38	38	51	51	51	76	76	76	76	76	76
20	160	**8**	**16**	23	**32**	**40**	40	53	53	53	**80**	80	80	80	80	80
21	168	**8**	15	**24**	28	34	42	**56**	56	56	56	84	84	84	84	84
22	176	**8**	**16**	22	29	35	44	44	59	59	59	**88**	88	88	88	88
23	184	**8**	15	23	31	37	46	46	61	61	61	61	92	92	92	92
24	192	**8**	**16**	**24**	**32**	38	**48**	48	**64**	64	64	64	**96**	96	96	96
25	200	**8**	15	22	29	**40**	40	50	50	67	67	67	67	100	100	100
26	208	**8**	**16**	23	30	35	42	52	52	69	69	69	69	**104**	104	104
27	216	**8**	15	**24**	31	36	43	54	54	**72**	72	72	72	72	108	108
28	224	**8**	**16**	22	**32**	37	45	**56**	56	56	75	75	75	75	**112**	112
29	232	**8**	15	23	29	39	46	46	58	58	77	77	77	77	77	116
30	240	**8**	**16**	**24**	30	**40**	**48**	48	60	60	80	80	80	80	80	**120**

This table is based upon the assumptions that the child care is 8 hours in length (TO) and that the full enrollment is present for the full 8 hours (TH). This is unlikely to ever occur but it gives us a reference point to measure adult child contact hours in the most efficient manner. Based upon the relationship between TO and TH based upon the algorithms, select from one of the formulae from the previous page (formulae 1 - 4) to determine how well the actual Relatively Weighted Contact Hours (RWCH) match with this table. If the RWCH exceed the respective RWCH in this table, then the facility would be over ratio on ACR standards, in other words, they would be overpopulated.

Figure 1: Contact Hour Diagram Paradigm and Schematic



The above diagram (Figure 1) depicts how the number of staff and children help to construct the contact hour formula. Depending on when the children arrive and leave could change the shape from a trapezoid to a rectangle or square or triangle. Please see the following potential density distributions which could impact these changes in the above contact hour diagram.

Potential Density Distributions
Taking into Account Number of Children, Staff, and Exposure Time

Here are some basic key relationships or elements related to the Contact Hour (CH) methodology.

- RWCH = ACR
- CH = GS = NC
- NC and CH are highly correlated
- ACR and GS are static, not dynamic
- CH makes them dynamic by making them 2-D by adding in Time (T)
- ΣACR = GS
- GS = total number of children NC
- ACR = children / adult

ACR = Adult Child Ratio, GS = Group Size, RWCH = Relatively Weighted Contact Hours, NC = Number of Children.

Possible Density Displays of Contact Hours (Horizontal Axis = Time (T); Vertical Axis = NC):

This density distribution should result in the lowest CH but probably not very likely to occur. Essentially what would happen is that full enrollment would be a single point which means that the last child arrives when the first child is leaving. Very unlikely but possible.

This density distribution is probably the most likely scenario when it comes to CH in which the children gradually, albeit rather steeply, arrive at the facility and also leave the facility gradually. They don't all show up at the same time nor leave at the same time. However, the arriving and leaving will be a rather close time frame.

This scenario is unlikely but is used as the reference point for CH because it provides the most efficient model. This is where all the children arrive and leave at the same time. Very unlikely, but I guess it could happen. The important element here is its efficiency in that all contact hours are covered, so although a lesser amount of CH is not as efficient it does demonstrate compliance with ACR and GS which is one of the purposes of CH. As the bottom two distributions will demonstrate, CHs above this level would either depict a program that is open for an extended time or where there are too many children present and the facility is out of compliance with GS and/or ACR.

This distribution would indicate that the facility is open for an extended time and exceeds the number of total CH as depicted in the reference square standard. Although not out of compliance with GS or ACR, this could become a determining factor when looking at the potential overall exposure of adults and children when we are concerned about the spread of an infectious diseases, such as what happened with COVID19. Are facilities that are high on a CH measurement more prone to the spread of infectious diseases?

This depiction clearly indicates a very high CH and non-compliance with ACR and GS. This is the reason for designing the CH methodology which was to determine these levels of regulatory compliance as its focus.

Section 3: Program Quality Indicators

This section provides the program quality indicators (PQI) which along with the previous section dealing with staff child ratios and group sizes constitutes the new Integrated Program Monitoring system: CCEE Heart Monitor (CCEEHM App). These PQI were validated in a study in the province of Saskatchewan (Fiene, 2024).

The PQI represents staffing, program, parental involvement and key interactional observation indicators drawn from key indicator studies from 1980 - 2020 involving quality rating and improvement systems (QRIS), professional development, and program quality initiative observational studies. These indicators provide the process quality within the context of the structural quality provided by the contact hour metric depicted in the previous section. Both the contact hour and these PQI are intended to be used in an integrated fashion and compliance should be measured on both domains. By doing this a picture of structural and process quality will be possible.

By utilizing this new integrated program monitoring system it will provide a cost effective and efficient system for jurisdictions around the world. These metrics are based upon research studies completed in the USA and Canada from 2020-2024 (Fiene, 2025a,b,c).

INDICATOR 1): Number of ECE III Educators (AA and BA Level ECE Educators)

AI will review staff records to determine the number of staff who have these credentials in early childhood education. Record the number of ECEs with the appropriate qualifications and divide them by the total number of ECEs to come up with a percent for the center.

How to Measure:

Go to a Staff Information Summary form to obtain the data for this item. Under Certification, look for the following: Certification Date and Certification Level (Highest ECE Level Certified). The certification date should be earlier than the date of the review and the actual level of the certification. In this case, we are interested in the number of (ECEIII's). Record the number of ECEIII working at least 65 hours/month. Then record the number of total teaching staff working at least 65 hours/month below as well. Teaching staff is defined as staff who have a responsibility for working with the children and the programming. Determine the percentage by dividing the total number of staff into the total number of ECEIII Certified teaching staff, ECEIII Certified teaching staff is the numerator, and the total number of teaching staff is the denominator (ECEIII/Total number of teaching staff x 100% = Percent).

Scoring for PQI 1:

The total number of ECEIII Certified teaching staff _____(1.1)

The total number of teaching staff _____(1.2)

Total ECEIII teaching staff divided by the total number of teaching staff _____

(%). Then based on the percentage, you can find the score of 1-4 as per the chart below.

Circle the Appropriate Level	***1 = 0 to 25%***	***2= 26 to 50%***	***3 = 51 to 75%***	***4 = 76 to 100%***

INDICATOR 2): Stimulating and Dynamic Environment

The criteria for measuring this are drawn from Play and Exploration Guides that should be present in all CCEE programs. The program should be child centered. Children are viewed as competent learners, and they have the freedom to access classroom materials independently without adult intervention. The children are provided with meaningful choices through activity/learning centers. There is evidence of the children's interests and their projects in the learning environment.

How to Measure:

Below is the checklist of items that should be present to assess if the environment is both stimulating and dynamic for the children. You will want to observe that the following items are occurring in the classroom first. If you do not actually observe it occurring, then check the program plan to find documentation that it normally occurs but you just did not observe today. The checklist items would be found in *Play and Exploration* foundational materials.

Quality Early Learning Environments (Please record all that you observe Y or N):

1. Co-teaching is evident. Y/N _____(2.1)
2. Children are viewed as competent learners & can access materials independently. Y/N ____(2.2)
3. Authentic and meaningful materials are used with children. Y/N _____(2.3)
4. Children are provided with meaningful choices. Y/N _____(2.4)
5. Children's work, art and photos are displayed respectfully. Y/N _____(2.5)
6. Family photos are displayed in the early learning program. Y/N _____(2.6)
7. Documentation of learning is displayed and discusses holistic development. Y/N _____(2.7)
8. Environment reflects the culture and beliefs of the children, families and staff. Y/N _____(2.8)
9. Variety of books & other print materials are available throughout the classroom Y/N _____(2.9)
10. A variety of writing materials are accessible to children most of the time. Y/N _____(2.10)
11. There is evidence of the children's interests & projects in the classroom. Y/N _____(2.11)

Scoring for PQI 2:

Total up the number of items where you recorded a "Y" above that you observed (curriculum or in classrooms), divide by 11 x 100% to come up with a percent and record here _____%. Then based on the percentage, you can find the score of 1-4 as per the chart below.

Circle the Appropriate Level	***1 = 0 to 25%***	***2= 26 to 50%***	***3 = 51 to 75%***	***4 = 76 to 100%***

INDICATOR 3): Developmentally Appropriate Curriculum Based on Assessments of Each Child

The key for this quality key indicator is that the program is following an individualized prescribed planning document when it comes to curriculum. It does not mean it is a canned program, in fact, it shouldn't if it is based upon the individual needs of each child's developmental assessment. The assessor will ask to see what is used to guide the curriculum. There should be a written document that clearly delineates the parameters of the philosophy, activities, guidance, and resources needed for the particular curricular approach. There should also be a developmental assessment which is clearly tied to the curriculum. The developmental assessment can be home-grown or a more standardized off-the-shelf type of assessment, the key being its ability to inform the various aspects of the curriculum. The purpose of the assessments is not to compare children but rather to compare the developmental progress of individual children as they experience the activities of the curriculum.

The following key elements should be present when assessing this quality indicator.

- 1) The program practices emergent curriculum, allowing the interests of the children to determine the learning content. The curriculum is informed by individual developmental assessments of each child in the respective classrooms.
- 2) The children and educators are co-learners in the exploration of projects.
- 3) Learning activities of the children are documented, displayed in the learning environment and used to plan further learning activities. This can be assessed developmentally.

How to Measure:

Take a sample of 10 individual children's records and consider the above three elements for EACH record. You should be asking yourself if there is a clear link between an assessment and the developmentally appropriate curriculum so that an individualized learning approach is being undertaken and each child's developmental needs are taken into consideration. These records could be formal, such as portfolios kept for each child or a more informal, anecdotal type of record keeping. The key is that there is a record that can be looked at. It is not adequate if the teacher says they do it from memory – it needs to be written down and documented.

Cross check the child's record to the actual curriculum. Record all the instances (Y's) in which this occurs. All three blocks need to be checked for each record (1-10).

Emergent Curriculum is Practiced (3.1)

1 Y/N	2 Y/N	3 Y/N	4 Y/N	5 Y/N	6 Y/N	7 Y/N	8 Y/N	9 Y/N	10 Y/N

Key Element 1 +

Children and Educators are Co-learners (3.2)

1 Y/N	2 Y/N	3 Y/N	4 Y/N	5 Y/N	6 Y/N	7 Y/N	8 Y/N	9 Y/N	10 Y/N

Key Element 2 +

Learning Activities are Documented and Displayed and Used to Plan Future Learning (3.3)

1 Y/N	2 Y/N	3 Y/N	4 Y/N	5 Y/N	6 Y/N	7 Y/N	8 Y/N	9 Y/N	10 Y/N

Key Element 3 +

All three key elements must have a Y to get an overall score of Y. If all three key elements have a Y for that individual record, then record Y in the corresponding block in the overall score.

1 Ys =	2 Ys =	3 Ys =	4 Ys =	5 Ys =	6 Ys =	7 Ys =	8 Ys =	9 Ys =	10 Ys =

= Total of All Three Key Elements (3.4)

Scoring for PQI 3:

The number of positive records (all Ys for all three elements) where there is a crosswalk from developmental assessment to curriculum _____

Percent of positive records (all Ys) (divide the number of positive records by 10 x 100%) _____%. Then based on the percentage, you can find the score of 1-4 as per the chart below.

Circle the Appropriate Level	*1 = 0 to 25%*	*2= 26 to 50%*	*3 = 51 to 75%*	*4 = 76 to 100%*

INDICATOR 4): Opportunities for Staff and Families to Get to Know Each Other

There should be activities both within the center as well as off site where staff and parents have opportunities to meet and greet each other. Communication with family members is documented and enables early childhood providers to assess the need for follow-up. Early childhood providers hold regular office hours when they are available to talk with family members either in person or by phone. Family members are encouraged to lead the conversation and to raise any questions or concerns.

How to Measure:

Look for the following 3 examples in policies developed by the program and determine if they have been carried out with families. It will be necessary to interview staff to complete this indicator if you do not find the three examples in policies:

1. The program provides communication, education, and informational materials & opportunities for families that are delivered in a way that meets their diverse needs. Y/N_____(4.1)
2. The program communicates with families using different modes of communication, and at least one mode promotes two-way communication. Y/N _____(4.2)
3. The program demonstrates respect and engages in ongoing two-way communication. The program respects each family's strengths, choices, & goals for their children. Y/N _____(4.3)

Scoring for PQI 4:

Record the number of Yes's (Y's): _____(Range: 0 – 3) (Divide by 3 x 100% = _____%). Then based on the percentage, you can find the score of 1-4 as per the chart below.

Circle the Appropriate Level	*1 = 0 to 25%*	*2= 26 to 50%*	*3 = 51 to 75%*	*4 = 76 to 100%*

INDICATOR 5): Families Receive Information on Their Child's Progress Regularly Using a

Formal Mechanism

Based upon Indicator #3 above, the information gleaned from the developmental assessments should be the focus of the report or parent conference. Parental feedback about the assessment and how it compares to their experiences at home would be an excellent comparison point. All these interactions should be done in a culturally and linguistically appropriate way representing the parents being served.

How to Measure:

Look for the following four examples in policies developed by the program and determine if they have been carried out with families. Record the number of reports completed or parent conferences over the past year. It will be necessary to interview staff to complete this indicator if you cannot determine from records that the conferences or reports were completed.

- 1) The program does have regularly scheduled (at least 2xs/year) parent conferences in which the children's developmental progress is discussed AND provides the family with a report of their child's developmental progress. Y/N _____(5.1) (Score 3 points). If "Yes" then go to Number 4. If "No", then go to numbers 2 and 3.
- 2) The program has regularly scheduled (at least 2xs/year) parent conferences in which the children's developmental progress is discussed, but it does not provide a report to the parents on their child's developmental progress. Y/N _____(5.2) (Score 2 points).
- 3) If the program does not have regularly scheduled (at least 2xs/year) parent conferences, does it provide the family with a report of their child's developmental progress. Y/N _____(5.3) (Score 1 point). Go to Number 4.
- 4) All these interactions are done in a culturally and linguistically appropriate way representing the parents being served. Y/N _____(5.4) (Score 1 point)

Scoring for PQI5:

Add up the total points based on the Ys; this will range from "0" to "4". The only way a program can receive a "4", is if a program has regularly scheduled parent conferences at least 2xs/year and provides the family with a report of their child's progress; and it is done in a culturally and linguistically appropriate way.

Record the number of points: _____(Range: 0 - 4)

Total Score for Part 1 = _____

PART 2 - OBSERVATIONS:

INDICATOR 6): Educators Encourage Children to Communicate (Preschool Class)

Assessors will need to observe this item when they do their classroom observations. Initially you can ask educators or the director how children are encouraged to communicate but in order to gather reliable and valid information regarding this question/standard, it needs to be observed in the various interactions between staff and children. Things to look for would be more back and forth conversations rather than one-way conversations where educators are telling children what to do. Look for opportunities where children can describe what they are doing, how they feel about what they are doing, and why they are doing particular activities. Educators expand upon children's conversation.

These opportunities can occur anywhere in the classroom or outside, such as in dramatic play, tabletop activities or on the playground. Materials should be present that encourage communication such as toy telephones, puppets, flannel boards, dolls and dramatic play props, small barns, fire stations, or dollhouses. These create a lot of conversation among children as they assume many different roles.
Children also talk when there is an interested person who listens to them. The staff in a high-quality early childhood classroom will use both activities and materials to encourage growth in communication skills.

How to Measure:

Observe the classroom for a minimum of 15 minutes. Once completed, consider where the classroom

falls based on the following scale;
Score the classroom a 1 if the following occur:

- No activities used by staff with children to encourage them to communicate, for example: non talking about drawings, dictating stories, sharing ideas at circle time, finger plays, singing songs. Y/N _____(6.1)
- Very few materials accessible that encourage children to communicate. Y/N _____(6.2)

Score the classroom a 2 if the following occur (If the classroom does not have all 3 indicators but has 2 of the indicators then score this item 1+):

- Some activities are used by staff w/children to encourage them to communicate. Y/N _____ (6.3)
- Some materials are accessible to encourage children to communicate. Y/N _____(6.4)
- Communication activities are generally appropriate for the children in the group. Y/N _____ (6.5)

Score the classroom a 3 if the following occur (If the classroom does not have both indicators but has one of the indicators then score this item 2+):

- Communication activities take place during both free play and group times, for example: child dictates story about painting; small group discusses trip to store. Y/N _____(6.6)
- Materials that encourage children to communicate are accessible in a variety of interest centers, for example: small figures and animals in block area; puppets and flannel board pieces in book area; toys for dramatic play outdoors or indoors. Y/N _____(6.7)

Score the classroom a 4 if the following occur (If the classroom does not have both indicators but has one of the indicators then score this item 3+):

- Staff balance listening and talking appropriately for age and abilities of children during communication activities, for example: leave time for children to respond; verbalize for child with limited communication skills. Y/N _____(6.9)
- Staff link children's spoken communication with written language, for example: write down what children dictate & read it back to them; help them write notes to parents. Y/N _____ (6.10)

Scoring for PQI 6:

Total up the number of "Y's" and record the appropriate level. In order for a classroom to receive a particular score, all "Y's" must be checked for the appropriate level (1 - 4) from above or partial credit given in order to obtain a "+". If there is a "+" please also mark it in the box.

Circle the Appropriate Level	***1***	***2***	***3***	***4***

INDICATOR 7): Infant Toddler Observation (if applicable) **(Infant Classroom)**

Conversations and questions should be used with all children, even young infants. Conversations using verbal and nonverbal turn-taking should be considered when scoring. Most conversations and questions initiated by infants will be nonverbal, such as widening of baby's eyes or waving arms and legs. Observe staff response to such nonverbal communication. For infants and toddlers, the responsibility for starting most conversations and asking questions belongs to the staff. As children become more able to initiate communication, staff should modify their approach in order to allow children to take on a greater role in initiating conversations and asking questions. Staff should provide answers to questions used by children if children cannot answer, and as children become more able to respond, questions should start to include those that the child can answer. If there was not an infant classroom, skip this Indicator and please note that here and on the summary score sheet by marking N/A: _____

How to Measure:

Observe the classroom for a minimum of 15 minutes. Once completed, consider where the classroom falls based on the following scale;

Score the classroom a 1 if the following occurs:

- Staff never initiate turn-taking conversations with children, for example: rarely encourage baby to babble back; simple back and forth exchanges with verbal children never observed. Y/N _____(7.1)
- Staff questions are often not appropriate for children, or no questions are asked, for example: too difficult to answer; carry a negative message. Y/N _____(7.2)
- Staff respond negatively when children can't answer questions, for example: "You should know this"; "You did not listen". Y/N _____(7.3)

Score the classroom a 2 if the following occurs (If the classroom does not have all 3 indicators but has 2 of the indicators then score this item 1+):

- Staff sometimes initiate conversations with children, for example: babble back and forth with baby; copy baby's sounds; respond to baby's crying with verbal response; have short back and forth toddler interactions. Y/N _____(7.4)
- Staff sometimes ask children appropriate questions and wait for the child to respond, for example: ask baby if she likes toy and pay attention as baby smiles; ask toddler what he is eating and wait for him to think of word. Y/N _____(7.5)
- Staff respond neutrally or positively to children who can't answer questions. Questions asked are sometimes meaningful to children, for example: child responds with interest; does not ignore staff questions. Y/N _____(7.6)

Score the classroom a 3 if the following occurs (If the classroom does not have all 4 indicators but has 2 or more of the indicators then score this item 2+):

- Staff initiate engaging conversations with children throughout the observation, for example: show enthusiasm; use tone that attracts child's attention. Y/N _____(7.7)
- Staff often personalize questions and/or conversations for individual children, for example: talk about children's families, preferences, interests; what they are playing with; what they did over weekend; child's mood; use child's name. Y/N _____(7.8)

- Staff often pay attention to children's questions, verbal or nonverbal, and answer in a satisfying manner for the child. Y/N _____(7.9)
- Staff ask questions in which children show interest in answering, for example: make the questions funny or mysterious; use attractive tone; meaningful and not too difficult to answer. Y/N _____(7.10)

Score the classroom a 4 if the following occurs (If the classroom does not have both indicators but has one of the indicators then score this item 3+):

- Staff frequently have turn taking conversations with children throughout the observations. Many appropriate questions are used throughout the observation, during both play and routines. Y/N _____(7.11)
- Staff ask children appropriate questions, wait a reasonable time for child response, and then answer if needed, for example: "Are you hungry? . . . Yes, you are!"; "Where's the ball? . . . These it is! You found the ball". Y/N _____(7.12)

Scoring for PQI 7:

Total up the number of "Y's" and record the appropriate level. For a classroom to receive a particular score, all "Y's" must be checked for the appropriate level (1 - 4) from above or partial credit given in order to obtain a "+".

Circle the Appropriate Level	***1***	***2***	***3***	***4***

INDICATOR 8): Educators Use Language to Develop Reasoning Skills (Preschool)

Assessors will need to observe very carefully as this standard can be difficult to determine because it is tying language and cognition together. Again, this opportunity can occur in any setting in or out of the classroom because it is the basis for problem solving through the use of language. Also look for educators redirecting children's conversations when appropriate. Staff should use language to talk about logical relationships using materials that stimulate reasoning. Through the use of materials, staff can demonstrate concepts such as same/different, classifying, sequencing, one-to-one correspondence, spatial relationships, and cause and effect.

How to Measure:

Observe the classroom for a minimum of 15 minutes. Once completed, consider where the classroom falls based on the following scale;

Score the classroom a 1 if the following occur:

- Staff do not talk with children about logical relationships, for example: ignore children's questions and curiosity about why things happen, do not call attention to sequence of daily events, differences and similarity in number, size, shape, cause and effect. Y/N _____(8.1)
- Concepts are introduced inappropriately, for example: concepts too difficult for age and abilities of children, inappropriate teaching methods used such as worksheets without any concrete experiences; teacher gives answers w/o helping children to figure things out. Y/N _____(8.2)

Score the classroom a 2 if the following occur (If the classroom does not have both indicators but has one of the indicators then score this item 1+):

- Staff sometimes talk about logical relationships or concepts, e.g.: explain that outside time comes after snacks, point out differences in sizes of blocks children use. Y/N _____(8.3)

- Some concepts are introduced appropriately for ages and abilities of children in group, using words and experiences, for example: guide children with questions and words to sort big and little blocks or to figure out why ice melts. Y/N _____(8.4)

Score the classroom a 3 if the following occur (If the classroom does not have both indicators but has one of the indicators then score this item 2+):

- Staff talk about logical relationships while children play with materials that stimulate reasoning, for example: sequence cards, same/different games, size and shape toys, sorting games, numbers and math games. Y/N _____(8.5)
- Children are encouraged to talk through or explain their reasoning when solving problems, for example: why they sorted objects into different groups, in what way two pictures are the same or different. Y/N _____(8.6)

Score the classroom a 4 if the following occur (If the classroom does not have both indicators but has one of the indicators then score this item 3+):

- Staff encourage children to reason throughout the day, using actual events and experiences as a basis for concept development, e.g.: children learn sequence by talking about their experiences in the daily routine or recalling the sequence of a cooking project. Y/N _____(8.7)
- Concepts are introduced based upon children's interests or needs to solve problems, for example: talk children through balancing a tall block building, help children figure out how many spoons are needed to set a table. Y/N _____(8.8)

Scoring for PQI 8:

Total up the number of "Y's" and record the appropriate level. In order for a classroom to receive a particular score, all "Y's" must be checked for the appropriate level (1 - 4) from above or partial credit given in order to obtain a "+".

Circle the Appropriate Level	***1***	***2***	***3***	***4***

For quality key indicators 9 and 10 it is recommended that these be assessed/observed throughout the observation period and not just during key activity times. These two quality key indicators should be observed in two-minute blocks over ten sequences for a total of 20 minutes. These two items should also be used with each age group being assessing.

INDICATOR 9): Educators Listen Attentively When Children Speak

This quality indicator focuses on the early childhood educator(s) looking directly at the children with nods, rephrasing their comments, and engaging in conversations. Children should have the undivided attention of the specific educator they are addressing. Educators should not be looking away or pre-occupied with others. They should be at the child's level making eye contact. The intent is to observe all children and educators in the room.

How to Measure:

Do this in timed 2-minute observations recording each time you observe this occurring. Record at least 10 different observation periods. These do not need to be consecutive in order to fully observe classrooms and educators. Please use the following scale to assess your recordings: Likert Scale (1-4) where 1 = Never/Not at All; 2 = Somewhat/Few Instances; 3 = Quite a Bit/Many Instances; 4 = Very Much/Consistently):
Make the actual recordings using the Likert Scale (1-4) above for each individual observation and record in each cell below.

09 Observations:

09.1	2	3	4	5	6	7	8	9	09.10

Scoring for PQI 9:

Once all the observations are made, add up the results from the Likert Scale (1-4) and record the total number here: _____(Range: 10 - 40)(Divide this result by 10) = _____(1-4)(Round upward or downward to the whole number (3.7 = 4; 2.2 = 2)).

Circle the Appropriate Level	***1***	***2***	***3***	***4***

INDICATOR 10): Educators Speak Warmly to Children

This quality indicator focuses on the early childhood educator(s) always engaging in a caring voice and body language with every child. Educators do not use harsh language or commands in speaking to children, but rather again are on the child's level making eye contact. Think of the way Fred Rogers would engage his audience where you always felt you were the most important person in the world when he talked to the TV.

How to Measure:

Do this in timed 2-minute observations recording each time you observe this occurring. Record at least 10 different observation periods. Please use the following scale to make your recordings: (This item is on a Likert Scale (1-4) where 1 = Never/Not at All; 2 = Somewhat/Few Instances; 3 = Quite a Bit/Many Instances; 4 = Very Much/Consistently):
Make the actual recordings using the Likert Scale (1-4) above for each individual observation and record in each cell below.
10 Observations:

10.1	2	3	4	5	6	7	8	9	10.10

Scoring for PQI 10:

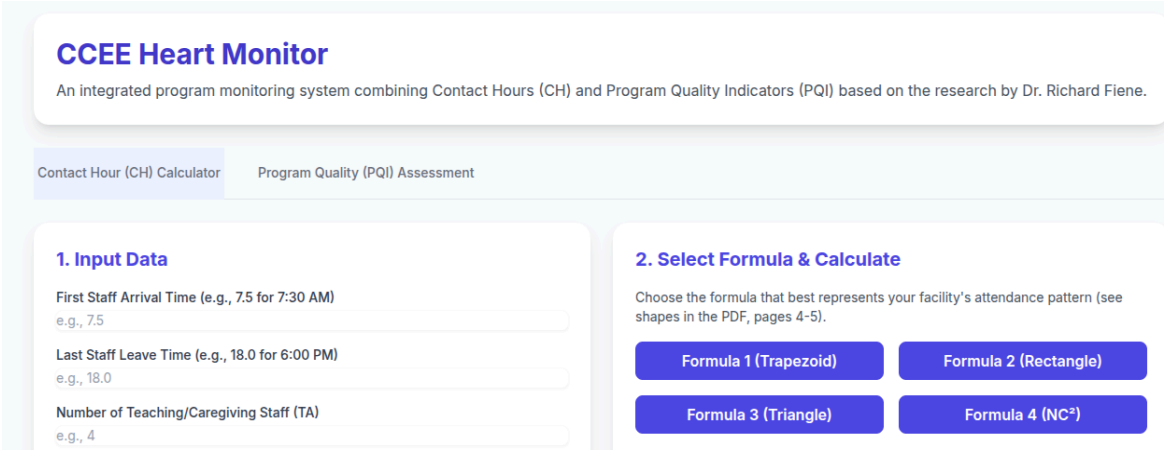
Once all the observations are made, add up the results from the Likert Scale (1-4) and record the total number here: _____(Range: 10 - 40) (Divide this result by 10) = _____(1-4). (Round upward or downward to the whole number (3.7 = 4; 2.2 = 2)).

Circle the Appropriate Level	***1***	***2***	***3***	***4***

Program Quality Indicators Artificial Intelligence (PQIAI) Scoring Protocol

LEVEL	Standardized Scores	Actual Scores
High Quality	**Mixed Age: 36+** **Preschool: 32+** **Infant-Toddler: 28+**	Mixed Age: _____ Preschool: _____ Infant-Toddler: _____
High - Mid Quality	**Mixed Age: 30 – 35** **Preschool: 26 - 31** **Infant-Toddler: 22 - 27**	Mixed Age: _____ Preschool: _____ Infant-Toddler: _____
Mid – Low Quality	**Mixed Age: 20 – 29** **Preschool: 16 - 25** **Infant-Toddler: 12 - 21**	Mixed Age: _____ Preschool: _____ Infant-Toddler: _____
Low Quality	**Mixed Ages: 19 or less** **Preschool: 15 or less** **Infant-Toddler: 11 or less**	Mixed Age: _____ Preschool: _____ Infant-Toddler: _____

Here is the opening screen to the Child Care and Early Education Heart Monitoring App (CCEEHM):


CCEE Heart Monitor
An integrated program monitoring system combining Contact Hours (CH) and Program Quality Indicators (PQI) based on the research by Dr. Richard Fiene.
Contact Hour (CH) Calculator
Program Quality (PQI) Assessment
1. Input Data
First Staff Arrival Time (e.g., 7.5 for 7:30 AM)
e.g., 7.5
Last Staff Leave Time (e.g., 18.0 for 6:00 PM)
e.g., 18.0
Number of Teaching/Caregiving Staff (TA)
e.g., 4
2. Select Formula & Calculate
Choose the formula that best represents your facility's attendance pattern (see shapes in the PDF, pages 4-5).
Formula 1 (Trapezoid)
Formula 2 (Rectangle)
Formula 3 (Triangle)
Formula 4 (NC²)

Here is the CCEE Heart Monitor Application: The ***Child Care and Early Education Integrated Program Monitoring System.*** It has two main sections, accessible through tabs:

1. **Contact Hour (CH) Calculator:** Input your facility's operational data to calculate the Contact Hour metric, which helps in analyzing structural quality. You can also include square footage for an expanded calculation.
2. **Program Quality (PQI) Assessment:** Go through the 10 indicators to evaluate the process quality of an early education program. The tool will automatically score each indicator and provide a final quality level based on the age group you select.

You can fill out the forms in each section and the application will compute the results for you in real-time. The tools that go along with these forms are appended to this document after the source code. You will need the tools for data collection and for interpreting the results from the Application so review these before opening the App. It will help familiarize you with the key data elements and the scoring system for this program monitoring systems approach.

The link to the CCEE Heart Monitor:

Alpha Version

https://g.co/gemini/share/f5397233272a

Beta Version

https://g.co/gemini/share/13dafe5005e9

References:

Fiene, R. (2025a). Potential Solution to the Child Care Trilemma Revisited: Finding the "Right Rules" - The Holy Grail of Early Care and Education, *Exchange*, Summer 2025, pps 64-66.

Fiene, R. (2025b). Finding the Rules that Work: An emerging paradigm promised to close the gap between regulatory compliance scores and quality of childcare services. *American Scientist, 113, 16-21, (January-February 2025).*

Fiene, R. (2025c). The Uncertainty–Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies. *Knowledge*, *5*(2), 8. https://doi.org/10.3390/knowledge5020008

Fiene, R. (2024). *Regulatory Compliance and Monitoring Systems*. Research Institute for Key Indicators and National Association for Regulatory Administration.

Freer, R. & Fiene, R. (2023). Regulatory Compliance and Quality Programs: Constraints and Opportunities for Integration. *Journal of Regulatory Science*, *11*(1), 1-10. https://doi.org/10.21423/JRS.REGSCI.111264

Fiene, R. (2022). Regulatory compliance monitoring paradigms and the relationship of regulatory compliance/licensing with program quality: A policy commentary. *Journal of Regulatory Science 10:1–7*.

Fiene, R. & Stevens, S. (2021). *Contact Hours as a New Metric Replacing Group Size and Staff-Child Ratios as well as a New Metric for COVID19 Thresholds.* Research Institute for Key Indicators/Penn State.

Fiene, R. (2019). A treatise on the theory of regulatory compliance*, Journal of Regulatory Science*, *Volume 7(1), pps 1-3.* (doi.org/10.21423/jrs-v07fiene).

Kontos, S. & Fiene, R. (1987). Child care quality, compliance with regulations, and children's development: The Pennsylvania Study, in *Quality in Child Care: What Does Research Tell Us?*, Phillips, editor, Washington, D.C.: National Association for the Education of Young Children.

Fiene, R. & Nixon, M. (1985). Instrument based program monitoring and the indicator checklist for child care, *Child Care Quarterly, 14*(3), 198-214.

NARA White Paper on Quality Indicators

Richard Fiene PhD

1. Overview

This paper introduces the concept and application of Quality Indicators in the Child Care and Early Education (CCEE) field. It begins by delineating how quality has historically been categorized into two dimensions: structural quality and process quality, as established in the research literature (Morgan, 1979; Peisner-Feinberg et al., 2001; Duncan, 2003). The paper then examines a range of models and systems to assess their effectiveness in achieving high-quality child care.

A central focus is the Theory of Regulatory Compliance, which offers a unifying framework for linking quality and regulatory compliance. This theory supports the development of a key indicator methodology that has been applied extensively across the United States and Canada (Fiene, 2019).

Building on this conceptual foundation, the paper highlights several systems that contribute to child care quality, including accreditation, professional development, training, technical assistance, Quality Rating and Improvement Systems (QRIS), and observational assessments such as the Environmental Rating Scales (ERS) (Fiene, 2022). Each of these systems is examined for its contribution to identifying key quality indicators.

Following this analysis, the paper introduces a consolidated Quality Indicators Tool that has evolved from these multiple systems (Fiene, 2024). The concluding section considers how this integrated approach to quality and compliance can be applied to other human service fields beyond early childhood care and education.

2. Introduction to the Theory and Methodology

Child care and early education (CCEE) quality has been defined in the research literature along a continuum between structural and process quality. Structural quality refers to countable, objective standards, while process quality reflects the more nuanced interpersonal dynamics between adults and children.

Structural quality is typically associated with licensing rules that safeguard children's health and safety in out-of-home care. Process quality, on the other hand, captures the nature of teacher-child interactions. These are often assessed through quality observation tools such as

the Environmental Rating Scales (ERS), commonly used within Quality Rating and Improvement Systems (QRIS) (Peisner-Feinberg et al., 2001; Duncan, 2003).

Process quality is often considered the core of early childhood quality. It captures the intricate details of what happens in classrooms during individual and group interactions. This includes positive and engaging language, emotional climate, and opportunities for children to problem-solve. Process quality emphasizes how well teachers facilitate these experiences, either through direct engagement or by arranging the learning environment to support them.

Structural quality refers to surrogate measures such as compliance with staff-child ratios, group sizes, or the number of regulatory violations. It does not typically assess emotional tone, instructional style, or classroom atmosphere. Occasionally, it may include aspects of curriculum, but its primary focus is on health and safety standards; those factors intended to prevent harm rather than promote enrichment. Enhancing the child's learning environment is primarily addressed through process quality.

From an organizational perspective, structural quality is most commonly found in licensing rules and regulations. Process quality, by contrast, is captured through tools such as the Environmental Rating Scales (ERS) or the Classroom Assessment Scoring System (CLASS). Although these tools were originally designed for standalone use, they are now more commonly embedded within broader quality initiatives. ERS is often used within QRIS, and CLASS is widely used in Head Start.

Structural and process quality are best understood as complementary. Structural quality provides a foundational layer, while process quality builds upon that base to support child development more holistically.

Another useful way to conceptualize quality is through the image of a spectrum. Structural and process quality can be placed at opposite ends of a continuum that reflects the range of quality interventions developed over the past 40 years. Interventions that align with structural quality include licensing, QRIS, Head Start Performance Standards, accreditation, and professional development systems. Process quality is primarily represented by tools such as ERS and CLASS.

This spectrum can be imagined as a prism: instead of separating light into colors, it separates the dimensions of quality into distinct intervention systems, each contributing in a different way to the overall goal of high-quality care.

The Original Model:

Gwen Morgan introduced the concept of a quality spectrum in a 1979 article published in Young Children (Morgan, 1979), which examined the components of child care quality. In that article, quality was categorized into two broad system types: regulatory and non-regulatory.

Regulatory components included licensing, contracting, best practices, credentialing, rate setting, and accreditation. Non-regulatory components included professional development and training systems, referral and resource agencies, advocacy organizations, and public education.

The Morgan Model was one of the earliest efforts to integrate these systems into a unified framework for understanding child care quality. Since its introduction, the model has expanded to include additional systems that have emerged over time, such as new accreditation frameworks, Quality Rating and Improvement Systems (QRIS), and statewide professional development and technical assistance initiatives—particularly those that use coaching or mentoring approaches.

The next section introduces a licensing and regulatory science framework that depicts the relationship between regulatory compliance and program quality.

Theory of Regulatory Compliance:

The Theory of Regulatory Compliance (TRC) (Fiene, 2019) marked a major breakthrough in how key indicators are conceptualized and applied in the licensing and monitoring of child care quality. It has been described as a paradigm shift that moved the licensing field and regulatory science away from a uniform monitoring approach toward a differential monitoring model (Fiene, 2025c,d). This transition enabled the development of the key indicator methodology, which made abbreviated inspections possible. Without TRC, the traditional model of uniform program monitoring might have remained the standard for conducting all child care inspections - but in practice, it proved insufficient.

Uniform monitoring was the dominant approach for licensing and oversight through the 1970s and 1980s. However, as evidence accumulated, it became clear that a more targeted method was needed. This shift laid the groundwork for the application of key indicators in a more nuanced and predictive way.

TRC serves as an overarching framework that explains how structural and process quality interact. One of its most important insights is the value of substantial compliance with structural quality rules. This was demonstrated through the identification of a ceiling effect when comparing structural and process quality across various systems, including licensing, Head Start, accreditation, and QRIS. Among these, licensing exhibited the most pronounced ceiling effect and, in some cases, a diminishing returns pattern when moving from substantial compliance to full 100 percent compliance. Although all structural systems showed similar tendencies, process quality followed a different pattern, typically a linear relationship with a normal data distribution. In contrast, structural quality data were non-linear and positively skewed. These contrasting patterns have been confirmed repeatedly in CCEE research over the past five decades.

As a result, TRC has helped reframe how licensing decisions are made. It recognizes that substantial compliance may be sufficient for issuing a full license, and in some cases, may be more predictive of program quality than full compliance. This perspective has also led to the use of abbreviated or targeted inspections that focus on key predictor rules and high-risk standards. This shift established the foundation for differential monitoring.

Importantly, differential monitoring allows quality indicators to be integrated into licensing and program oversight, something that was not easily achieved under uniform monitoring models. The structure and methodology of differential monitoring are described in the following section.

Differential Monitoring and the Key Indicator Methodology:

Recent policy studies (Freer & Fiene, 2023) confirm that quality indicators can be systematically identified and integrated into the licensing and regulatory framework. These indicators form the foundation of the differential monitoring model, serving as anchors for both structural and process quality.

Key indicators have repeatedly been shown to statistically predict overall compliance with the full set of rules, regulations, and standards through both ongoing research and analysis of key indicator system effectiveness in states and provinces that have incorporated key indicator systems into their licensing programs. This predictive relationship has been repeatedly validated across systems such as licensing, QRIS, Head Start, accreditation, and ERS. It also supports the development of a new quality indicator scale that is based on the same statistical methodology but is applied to nonregulatory standards that measure quality in early childhood development and education using standardized quality measurement tools.

Originally developed for child care licensing (Fiene & Nixon, 1985), the key indicator methodology has since been successfully applied to the identification of key compliance indicators, key risk indicators, key performance indicators, and key quality indicators. It is essential that any use of this methodology follows the original framework outlined in Fiene's research and NARA's implementation guidelines.

Two methodological components are particularly critical: the weighting of rules and the dichotomization of data. These techniques represent intentional departures from conventional predictive models in the testing and measurement literature. When properly applied, they help mitigate or eliminate false positives and false negatives in licensing decisions, particularly when determining whether a facility should receive a license (Fiene, 2024).

From a statistical perspective, correlations between structural and process quality have occasionally been significant, but they are generally modest. This is largely due to differences in

how the two types of data are structured and measured. Structural quality data tend to show a ceiling effect and a positively skewed distribution, whereas process quality typically follows a more linear relationship and exhibits a normal distribution. These distinct statistical profiles help explain why correlations between structural and process quality are often weak, even though each provides meaningful insight on its own.

A key factor underlying this divergence is the binary nature of structural quality measurement. Licensing rules are inherently dichotomous: a provider either complies with a given rule or does not. This yes/no structure limits the ability of structural data to capture variation in performance above the minimum threshold. Once compliance is achieved, there is no higher score available under the traditional model.

In contrast, process quality assessments use ordinal or interval scales to capture a broader continuum of performance from very low quality to exemplary practice. These tools can account for gradations in quality, such as partially met expectations, developing skills, or consistently high engagement. As a result, process quality tools are better equipped to differentiate among programs performing above the basic compliance level.

This difference in measurement design not only contributes to the ceiling effect in structural quality but also restricts its ability to meaningfully correlate with process quality across the full range of provider performance. Structural measures tend to group most providers at or near the top, making it difficult to distinguish between high and moderate performers. By contrast, process quality ratings offer a more nuanced and scalable view of practice.

Despite these limitations, both structural and process quality are generally effective at identifying the lowest-performing programs. Structural violations and poor-quality interactions tend to co-occur in these settings, reinforcing the value of using both types of data to flag high-risk environments

The Emergence of the Regulatory Compliance Scale:

Challenges in differentiating high-performing and average-performing programs prompted the development of a new structural quality metric: the Regulatory Compliance Scale (RCS) (Fiene, 2025b). This scale was introduced for two primary reasons. First, the RCS aligns with the Theory of Regulatory Compliance by emphasizing substantial compliance and incorporating a categorical scoring structure. Second, its ordinal format corresponds well with widely used process quality tools, which typically apply a 1–7 rating scale.

By establishing a categorical structure, the RCS enables structural quality to be analyzed on more equal footing with process quality from a statistical measurement standpoint. Pilot testing in several jurisdictions suggests that the RCS is a more effective comparative tool than

traditional methods that rely solely on violation counts or frequency data drawn from rule, regulation, or standard compliance (Fiene, 2024).

While these ideas have been addressed individually in prior literature (Trivedi, 2015), this paper brings them together to demonstrate the broader impact of the Theory of Regulatory Compliance on both structural and process quality assessment. Future replication of these findings by research psychologists and regulatory scientists would support an important policy advancement: recognizing substantial compliance as a sufficient threshold for full licensure and embedding differential monitoring as a standard regulatory practice across the CCEE field.

To support such a shift, the ceiling effect associated with structural quality must be reliably replicated when compared to the more evenly distributed nature of process quality scores (Fiene, 2025d).

3. The Various Quality Systems: Accreditation, Professional Development, Training, & Technical Assistance, Quality Observations, and Quality Rating & Improvement Systems (QRIS), and the Development of a Quality Indicators Scale

The same methodology used to identify Key Compliance Indicators has also been applied to the identification of Key Quality Indicators, which are the focus of this section. This Key Indicator Methodology is adaptable to any system governed by rules, regulations, or standards. Its effectiveness has been demonstrated in child welfare (Fiene, 1987), foster care reviews (Stevens, Fiene, Blevins, & Salzer, 2020), and in the identification of Key Performance and Key Risk Indicators for the Head Start Monitoring System.

The first quality initiative examined is accreditation, which has evolved significantly in the child care and early education field. Some accreditation systems are based on expert consensus or literature review, while others rely on empirically validated research methods. This section focuses specifically on one accreditation system developed using the key indicator approach discussed earlier (Fiene, 1995, 1996). The National Early Childhood Program Accreditation (NECPA) system was developed and field-tested in the early 1990s as a cost-effective, efficient alternative to existing models. Its standards overlapped with both the National Association for the Education of Young Children's (NAEYC) accreditation framework and the newly developed Caring for Our Children standards (AAP & APHA, 1992). In validation studies (Fiene, 1996), NECPA was found to be highly correlated with the more prominent NAEYC system.

The second major initiative involves professional development, training, and technical assistance systems, which have been implemented across all states using quality incentive funds from the Child Care Development Block Grant. Application of the key indicator methodology found that coaching and mentoring were more effective than traditional methods in improving

teacher-child interactions (Fiene, 2002). As a result, many states have shifted from workshop-based training to a blended model emphasizing coaching and mentoring.

The third major quality initiative is the implementation of Quality Rating and Improvement Systems (QRIS) as a supplement to traditional licensing systems. QRIS has proven effective in improving both structural and, in many cases, process quality across participating states (Zellman & Fiene, 2012). Approximately half of all states have adopted QRIS as of this writing. Within QRIS frameworks, Key Quality Indicators (KQIs) have been developed, particularly in the domains of family engagement and communication (Fiene, 2014). These indicators are central to program scoring within QRIS systems. Programs that meet KQIs are statistically more likely to receive higher QRIS ratings, typically at Level 3 or 4.

The final initiative discussed is quality observation, typically conducted using Environmental Rating Scales (ERS) (Harms, Clifford, & Cryer, 2012). ERS tools are widely used in QRIS systems as well as in other quality measurement contexts. A research study was conducted to determine whether ERS data could yield identifiable Key Quality Indicators. The study found that subscales related to language exchange and reasoning skills between teachers and children served as strong predictors of overall ERS scores.

This leads to the current state of Key Quality Indicator development. Most recently, a new Key Quality Indicators tool and software application has been proposed (Fiene, 2025a). The tool integrates findings from the initiatives described above, combining structural and process quality metrics into a single platform intended to improve both efficiency and comprehensiveness. However, the tool is currently in beta testing and requires further empirical validation to determine its long-term utility in the CCEE field. If validated, it may reshape how structural and process quality are assessed: consolidating currently separate systems into a unified approach.

4. Conclusion and Expansion Beyond Child Care

Quality indicators have expanded the scope of program monitoring, moving the field from uniform monitoring to differential monitoring, and now toward integrated monitoring approaches that explicitly incorporate quality (Freer & Fiene, 2023). Historically, key indicator methodologies focused primarily on regulatory compliance, with quality dimensions often excluded from formal review processes. This has changed with the identification of Key Quality Indicators (KQIs) drawn from accreditation systems, Quality Rating and Improvement Systems (QRIS), professional development and technical assistance initiatives, and observational tools such as the Environmental Rating Scales (ERS).

The resulting integrated monitoring model offers jurisdictions a promising framework for evaluating both structural and process quality in their inspection systems. It represents an opportunity to align compliance monitoring with broader goals related to service effectiveness and developmental outcomes, particularly in early childhood settings.

This paper aims to provide guidance on the research supporting KQIs and how these indicators can be used to predict overall quality across structural and process dimensions. While the examples are drawn from the Child Care and Early Education (CCEE) field, the underlying approach and methodology are applicable across a wide range of human service systems, including child residential and adult residential programs. Indeed, this methodology can be applied in any setting governed by rules, regulations, or standards as discussed earlier.

However, there are limitations to expanding the model into other human services contexts, particularly adult care. One significant challenge is the absence of established population-wide quality evaluation systems for older youth and adult service recipients. While evaluation methods do exist in these domains, they are often designed around the experience of the individual, assessing the quality of care received by a specific person rather than the overall performance of the service provider. This distinction complicates the development of system-wide Key Quality Indicators.

Additionally, it is inherently easier to define and measure growth and development in early childhood due to well-established developmental and educational milestones. In contrast, defining "quality" for adolescents, adults, or aging populations can be more subjective and variable. What constitutes meaningful progress or enrichment later in life is harder to quantify, and normative benchmarks are less universal.

Key Indicator Methodology itself is also relatively rare in adult care regulatory systems. This may be due in part to the historical concentration of quality indicator research and application within child care, as well as the more mature and standardized nature of child care regulation compared to systems serving adults. As a result, further research and development will likely be needed to adapt and expand the key indicator framework for adult populations. This may include the creation of tiered, provider-level quality assessment tools that go beyond individual outcomes to measure systemic performance.

If such tools can be developed and validated, the integration of quality indicators into adult care licensing systems could mirror the transformative impact seen in child care, enabling targeted monitoring, data-informed licensing decisions, and a stronger link between compliance and meaningful quality outcomes.

References and Future Readings:

American Academy of Pediatrics and American Public Health Association (1992). *Caring for Our Children*, AAP & APHA, Washington, D.C.

Fiene, R. & Nixon, M. (1985). Instrument based program monitoring and the indicator checklist for child care, *Child Care Quarterly, 14(3),* 198-214.

Fiene & McDonald (1987). *Instrument based program monitoring for child welfare*, Portland, Maine: University of Southern Maine.

Fiene (1995). *National early childhood program accreditation standards*. Atlanta, Georgia: National Early Childhood Program Accreditation Commission.

Fiene (1996). Using a statistical-indicator methodology for accreditation, in *NAEYC Accreditation: A Decade of Learning and the Years Ahead*, S. Bredekamp & B. Willer, editors, Washington, D.C.: National Association for the Education of Young Children.

Fiene (2002). Improving child care quality through an infant caregiver mentoring project, *Child and Youth Care Forum, 31*(2), 75-83.

Fiene, R. and Kroh, K.E. (2016). Human Services Licensing Measurement, Regulatory Compliance, and Program Monitoring Systems Chapter 11, part of *NARA Licensing Curriculum*, Research Institute for Key Indicators & the National Association for Regulatory Administration: Minneapolis, MN, USA.

Fiene, R. (2019). A treatise on regulatory compliance. *Journal of Regulatory Science* 7:1–3.

Fiene, R. (2022). Regulatory compliance monitoring paradigms and the relationship of regulatory compliance/licensing with program quality: A policy commentary. *Journal of Regulatory Science,* 10:1–7.

Fiene, R. (2024). *Regulatory compliance and monitoring systems ebook text*. National Association for Regulatory Administration licensing curriculum chapter on licensing measurement and program monitoring systems. Published online www.naralicensing.org/key-indicators.

Fiene, R. (2025a). The Child Care and Early Education Heart Monitor, submitted for publication in the *International Critical Childhood Policy Journal.*

Fiene, R. (2025b). The Development of a Regulatory Compliance Scale, *Scholarly Community Encyclopedia,* June, 2025, 1-18.

Fiene, R. (2025c). Finding the Rules that Work: An emerging paradigm promised to close the gap between regulatory compliance scores and quality of childcare services. *American Scientist, 113,* 16–21.

Fiene, R. (2025d). Potential Solution to the Child Care Trilemma Revisited: Finding the "Right Rules" - The Holy Grail of Early Care and Education, *Exchange*, Summer 2025, pps 64-66.

Fiene, R. (2025e). The Uncertainty–Certainty Matrix for Licensing Decision Making, Validation, Reliability, and Differential Monitoring Studies. *Knowledge, 5(2), 8.* https://doi.org/10.3390/knowledge5020008

Freer, R. & Fiene, R. (2023). Regulatory Compliance and Quality Programs: Constraints and Opportunities for Integration. *Journal of Regulatory Science, 11(1)*, 1-10. https://doi.org/10.21423/JRS.REGSCI.111264.

Harms, T., Clifford, R., & Cryer, D. (2012). *Early Childhood Environmental Rating Scale*. Columbia University Press, New York, New York.

Kontos, S. & Fiene, R. (1987). Child care quality, compliance with regulations, and children's development: The Pennsylvania Study, in *Quality in Child Care: What Does Research Tell Us?*, Phillips, editor, Washington, D.C.: National Association for the Education of Young Children.

Lahti, M.; Elicker, J.; Zellman, G.; Fiene, R. (2014). Approaches to validating child care quality rating and improvement systems (QRIS): Results from two states with similar QRIS type designs. *Early Childhood Research Quarterly, 30*, 280–290.

Maxwell, K.; Miranda, B.; Harris, P.; Johnson, N.; Padilla, C.; Banghart, P.; Verhoye, A.; Early, D.M. (2024). *Child care and early education licensing literature review.* In OPRE Report #2024-199; Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services: Washington, DC, USA.

Miranda, B.; Ekyalongo, Y.; Franchett, A.; Maxwell, K. (2022). *Monitoring Practices Used in Child Care and Early Education Licensing.* In OPRE Report #2022-137; Office of Planning, Research, and Evaluation, Administration for Children and Families, Department of Health and Human Service: Washington, DC, USA.

Morgan, G. (1979). Methods for Achieving Quality Child Care, *Young Children, Volume 34, Number 6,* pps 22-27.

National Institute of Child Health and Human Development Early Child Care Research Network and Duncan, G.J. (2003), Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development. Child Development, 74: 1454-1475. https://doi.org/10.1111/1467-8624.00617

Payne, A.L. (2011). *Strong Licensing: The Foundation for a Quality Early Care and Education System*, National Association for Regulatory Administration: Minneapolis, MN, USA.

Peisner-Feinberg, E.S., Burchinal, M.R., Clifford, R.M., Culkin, M.L., Howes, C., Kagan, S.L. and Yazejian, N. (2001), The Relation of Preschool Child-Care Quality to Children's Cognitive and Social Developmental Trajectories through Second Grade. *Child Development, 72*: 1534-1553. https://doi.org/10.1111/1467-8624.00364

Stevens, Fiene, Blevins, Salzer (2020). Identifying Predictive Indicators: The State of Washington Foster Care Home Study, *Journal of Children and Youth Services*.

Trivedi, P.A. (2015). *Innovation in Monitoring in Early Care and Education: Options for States*, Office of the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services: Washington, DC, USA.

Zellman, G. and Fiene, R. (2012). *Validation of quality rating and improvement systems for early care and education and school-age care.* In OPRE and Child Trends, Office of Planning Research and Evaluation Administration for Children and Families, U.S. Department of Health and Human Services: Washington, DC, USA.

About the Author:

Dr Richard Fiene, a research psychologist, has spent his professional career in improving the quality of child care in various states, nationally, and internationally. He has done extensive research and publishing on the key components in improving child care quality through an Early Childhood Program Quality Indicator Model (ECPQIM) of training, technical assistance, quality rating & improvement systems, professional development, mentoring/coaching, regulatory science, licensing, risk assessment, differential program monitoring, key indicators, and accreditation. His research has also made significant contributions in regulatory science related to measurement and monitoring systems, such as instrument-based program monitoring, differential monitoring, key indicator methodology for compliance and quality, and risk assessment methodology. In prevention science, his research has led to the identification of key Regulatory indicators that keep children healthy and safe while in out of home child care settings.

Dr Fiene is a Professor of Psychology (ret) (Penn State University) and founding director of the Capital Area Early Childhood Research and Training Institute. He is presently a Research Psychologist and Regulatory Prevention Scientist for the Research Institute for Key Indicators, an affiliated data laboratory with the Edna Bennett Pierce Prevention Research Center at the Pennsylvania State University.

Dr Fiene is regarded as a leading international researcher/scholar on human services

licensing measurement and differential monitoring systems. His regulatory compliance law of diminishing returns has altered human services regulatory science and licensing measurement dramatically in thinking about how best to monitor and assess licensing rules and regulations through targeted and abbreviated inspections. The theory has also led to the issuing of human service licenses based on substantial regulatory compliance with all rules rather than full 100% regulatory compliance with all rules. This was a basic licensing and public policy paradigm shift which has impacted regulatory administration.

His research has led to the following developments: identification of herding behavior of two year olds, spatial acquisition device in young children & four states of space, national early care and education quality indicators, mathematical model (Contact Hours) for determining adult child ratio compliance, solution to the trilemma (quality, affordability, and accessibility) in child care delivery services, Stepping Stones to Caring for Our Children, NECPA: National Early Childhood Program Accreditation, online coaching as a targeted and individualized learning platform, validation framework for early childhood licensing systems and quality rating & improvement systems, an Early Childhood Program Quality Improvement & Indicator Model for better public policy decision making, Caring for Our Children Basics, Abbreviated Program Monitoring Inspections, Validation Framework for Licensing, Generic Key Indicator Rules, Regulatory Compliance Scoring Scale, RegalMetrics, and has led to the development of statistical techniques for dealing with highly skewed, non-parametric data distributions in human services licensing and regulatory systems, such as data dichotomization.

Dr Fiene had a long career in academia and governmental service. He was a research psychologist and regulatory scientist during his tenure with the Commonwealth of Pennsylvania's Office of Children, Youth, and Families and the Office of Licensing and Regulatory Administration where he was the research director for both offices. In academia he was a professor of psychology and human development at both the University of North Carolina and the Pennsylvania State University. At Penn State Harrisburg he was Department Head for both the psychology and human development programs during his tenure at the university.

At the national and international levels, Dr Fiene has been a senior research consultant to the National Association for Regulatory Administration, the Federal Office of Child Care, the Administration for Children and Families, and the Federal Department of Health and Human Services. His research has been disseminated to all 50 states and over 120 countries. In 2019, he was elected to the Early Childhood Exchange Leadership Initiative. He received the 2020 Distinguished Career Award from the Pennsylvania Association for the Education of Young Children. In 2023, his Key Indicator methodology for quality indicators received a Recognized Project of the Child Impact Initiative of the World Forum Foundation. Dr Fiene remains active in the regulatory prevention science and early childhood fields through the Edna Bennett Pierce Prevention Research Center at Penn State where he remains an affiliated faculty and a senior research psychologist. He has been a member of the American Psychological Society.